Minimax Optimality in Robust Detection of a Disorder Time in Poisson Rate

Nicole El Karoui, Stéphane Loisel, Yahia Salhi

Minimax Optimality in Robust Detection of a Disorder Time in Poisson Rate

Nicole El Karoui, Stéphane Loisel,
and Yahia Salhi

Abstract.
We consider the minimax quickest detection problem of an unobservable time of change in the rate of an inhomogeneous Poisson process. We seek a stopping rule that minimizes the robust Lorden criterion, formulated in terms of the number of events until detection, both for the worst-case delay and the false alarm constraint. In the Wiener case, such a problem has been solved using the so-called cumulative sums (cusum) strategy by Shiryaev [33, 35], or Moustakides [24] among others. In our setting, we derive the exact optimality of the cusum stopping rule by using finite variation calculus and elementary martingale properties to characterize the performance functions of the cusum stopping rule in terms of scale functions. These are solutions of some delayed differential equations that we solve elementarily. The case of detecting a decrease in the intensity is easy to study because the performance functions are continuous. In the case of an increase where the performance functions are not continuous, martingale properties require using a discontinuous local time. Nevertheless, from an identity relating the scale functions, the optimality of the cusum rule still holds. Finally, some numerical illustration are provided.

AMS 2000 subject classifications: Primary 62L15; Secondary 97M30; .
Keywords: Change-Point, Robust Sequential Detection, Poisson Process.

1. Introduction. In the Poisson quickest detection problem, one observes the jumps of an inhomogeneous counting process whose intensity suddenly changes at some unobservable disorder time, but whose intensity is "stable" beforehand and afterhand in some sense. As pointed in the introduction of Basseville and Nikiforov (1993) [6] *"It should be clear that abrupt changes by no means imply changes with large magnitude. Many change detection problems are concerned with the detection of small changes."*. The process being sequentially observed, the problem is then to detect the change-point as quickly as possible after it happens.

The classical fields of applications of detection problems include, among others, queueing theory, survival analysis and reliability [6]. Our main practical motivation was the fast detection of the onset of mortality shifts, where a proportional relationship between two mortality intensities, e.g. the insured mortality intensity and a reference, is assumed. A sudden change in the proportional relationship can induce serious financial consequences, and it is necessary to react as soon as data would suggest, see Barrieu et al. (2012) [5]. Similar problem also arise in prompt detection of shifts in insurance claims arrival. Note that in these examples, in general, *no information* is known about the distribution of the date of change. Therefore, in this paper, we consider a non-Bayesian setting in which the change-point is unknown but deterministic, in the spirit of the early papers of Page (1954) [25] and Lorden (1971) [20]. This framework, based on the so-called Lorden

procedure formulated as a minimax problem, belongs to the family of *robust* optimization, popular in statistical learning, see for e.g. Hastie et al. (2009) [15].
The minimax robust detection problem in continuous times has gained a renewed interest since the 90's. For example, the problem of detecting an abrupt change on the drift of a Wiener process is well understood. In particular, the cumulative sums strategy [25] (cusum for short) has been shown to be optimal with regard to the Lorden procedure, see Shiryaev (1996) [34], Beibel (1996) [8] and Moustakides (2004) [24]. Shiryaev (2009) [35] shed an interesting light on the history of this problem, and different developments depending on the fields of application.

In the classical sequential test analysis between continuous times processes, the statistic is the usual Log Sequential Probability Ratio process (LSPR-process) between the reference probability $\mathbb{P}$ (null assumption) and the alternative assumption $\widetilde{\mathbb{P}}$ (H1 assumption), see Dvoretzky, Kiefer and Wolfowitz (1953) [13] for the Poisson process. In the minimax detection problem, the test is based in the cusum strategy, consisting on sounding an alarm as soon as the LSPR-process related to $\mathbb{P}$ (no change) and $\widetilde{\mathbb{P}}$ (immediate change), *reflected at its maximum or at its minimum,* hits a barrier $m > 0$. The optimality relies on the characterization of the optimal detection time through its performance, i.e. the time until detection and the false alarm frequency, which are of critical importance. In the Wiener case, the performance functions have a very simple and universal form, which is a key element in the derivation of the optimality [24]. In this paper, we are interested in the extension of this optimality result to the cusum strategies in the Poisson case, by overcoming difficulties due to jumps and to the complex form of the performance functions, strongly related with the scale functions introduced in Lévy process theory, e.g. Bertoin (1998) [9], Pistorius (2004) [28], and reference therein. Note that, in discrete time, only asymptotic optimality result was shown by Mei et al. (2011.) [21].

It is worth mentioning that the Poisson disorder problem has been widely studied in the Bayesian setting, see Bayraktar et al. (2005) [7] and Chapter 5 of Poor and Hadjiliadis (2009) [29] and the reference therein. The optimal detection rule is well studied for different formulation but is very sensitive to the a priori distribution of the change time.

The remainder is organized as follows. In Section 2, we introduce the robust optimization problem, with a discussion on the criterium to be associated with this minimax problem. This discussion is followed by a rigorous presentation of the basic tools in counting process framework. In Section 3, as in sequential test analysis, we introduce the log-sequential probability ratio process (LSPR in short) and highlight an interesting link with the surplus process in ruin theory or with the equity process in insurance. Time rescaling procedure allows us to compare the distributions of the LSPR process under the probability measures $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ through their realization on the same probability space. In Section 4, we are concerned with the cusum processes, which are the log-sequential probability ratio process reflected at its running maximum or its running minimum, and show that some solutions of differential equations driven by the counting process appear. We also define the performance functions associated with the hitting time by the cusum processes of a barrier m making the distinction between an increase and a decrease in the intensity. In Section 5, we use differential calculus with jumps to solve a classical delay equation whose solutions may be proportional to some ruin distribution functions, or scale functions in Lévy's theory. Stochastic differential calculus is used to extend Itô's formula to a.e. differentiable monotonic functions with one jump, thanks to a discontinuous local time. We study the performance functions associated with positive barrier with the help of the scale functions and their primitives, and prepare the proof of their optimality by the introduction of new, well-suited martingales. Section 6 is dedicated to the proof of optimality of the cusum procedure with a given false alarm constraint. The proof relies on a modified criterion following ideas form Shiryaev (1996) [34] and Moustakides (1986, 2002) [22, 23]. To the best of our knowledge, it is the first optimality result in the Poisson case, for the minimax problem. Finally, in Section 7 we discuss some numerical methods suited to solve the delayed equation associated with the scale and performance functions, and overcome numerical instability problems.

2. Problem Formulation.

2.1. *General consideration.* Let N be a counting process with arrival times $(T_i)_{i\geq 1}$ and (possibly stochastic) intensity process $\check{\lambda} = (\check{\lambda}_t)_{t\geq 0}$. We are interested in detecting any disorder that can affect the dynamics of such a counting process.
(i) There are different ways to model such a change which may occur at some unobservable, non random time $\theta \in [0,\infty)$. The best suited (and more frequently used) to the detection problem is to keep the observed process N under the reference (or nominal) probability measure $\mathbb{P}$ associated with the intensity process denoted (λ_t) when the change never happenes, and to change the likelihood of occurrences by using an equivalent probability measure $\mathbb{P}^\theta$ characterized by its *likelihood ratio* $\mathcal{E}^\theta_{0,T}$ with respect to $\mathbb{P}$. Since the change in the intensity only occurs at time θ, we have $\mathcal{E}^\theta_{0,t} = 1$ for $t \leq \theta$. Therefore, we suppose that the intensity $\check{\lambda}$ undergoes a sudden change from the value λ_θ to the value $\rho\lambda_\theta$, where ρ is a known positive constant number different from 1. This formulation corresponds to the definition of a simplified Cox (1972) [10] proportional model framework, which is widely used in life insurance and survival analysis. Formally, we consider that,

$$\check{\lambda}_t = \lambda_t \mathbf{1}_{\{t<\theta\}} + \rho\lambda_t \mathbf{1}_{\{t\geq\theta\}}, \quad \rho \neq 1, \tag{1}$$

and denote $\Lambda_t = \int_0^t \lambda_s\, ds$ the cumulative intensity process when no change occurs.
(ii) In sequential theory, the strategies are based on the information acquired over time. As usual, it is modeled via a filtration $(\mathcal{F}_t)$ for which the process N and its intensity $\check{\lambda}$ are adapted. The sequential conditional probability ratio (SCPR on short) between t and T is given by $\mathcal{E}^\theta_{t,T} = \mathcal{E}^\theta_{0,T}/\mathcal{E}^\theta_{0,t}$ and plays a key role both in sequential testing theory and in quickest detection problems, but we are more concerned with quickest detection procedures [6].

2.2. *Robust detection problem.* Our main objective is to find a stopping rule T based on the filtration $\mathbb{F}$ to optimally raise an alarm for the breakpoint, with no *a priori* information on θ. The first step is to define a measure that quantifies its performance.
(i) For a robust estimation, Lorden [20] procedure advocates penalizing the detection delay via its worst-case value

$$C_{\mathrm{Lor}}(T) = \sup_{\theta\in[0,\infty]} \operatorname*{ess\,sup}_{\omega} \mathbb{E}^\theta\big[(T-\theta)^+ \big| \mathcal{F}_\theta\big], \tag{2}$$

where the ess sup takes the "worst possible observed date before the change" in the sense of providing no information on the true change as explained in Mei et al. [21][1]. In addition, the *false alarm* is monitored using the average run length to false alarm given by $\mathbb{E}^\infty[T] = \mathbb{E}[T]$. With these metrics, the quickest detection procedure reduces to solve the following minimax optimization problem:

$$\inf_T \{C_{\mathrm{Lor}}(T) |\, \mathbb{E}(T) \geq \pi\}, \quad \text{where } \pi \text{ is a given threshold.} \tag{3}$$

[1]Mathematically, it is the smallest constant number y_θ such that $y_\theta \geq \mathbb{E}^\theta\left[(T-\theta)^+ \big| \mathcal{F}_\theta\right]$, $\mathbb{P}^\theta$-a.s.

This framework serves to design the optimal stopping rule T^* with the desired characteristics based on a linear delay criterion, well-adapted to constant intensities. Thus, since we are working on (possibly) random intensity process, we are looking for a criterium which is robust with respect to *time rescaling* procedure, both in the criterion and the false alarm constraint, as the number of events until detection. More precisely, we consider the following criterion, and the minimax optimization problem,

$$\inf_T \{C(T)|\, \mathbb{E}(N_T) \geq \pi\}, \quad \text{where} \quad C(T) = \sup_{\theta\in[0,\infty]} \operatorname*{ess\,sup}_{\omega} \mathbb{E}^\theta\big[(N_T - N_\theta)^+\big|\mathcal{F}_\theta\big]. \tag{4}$$

In the same vein, Moustakides [24] has initiated a modification of the Lorden criterion that replaces expected delays with Kullback-Leibler divergences.

2.3. *Basic framework.* We give a precise mathematical framework for the different notions introduced above, concerning the counting process with varying intensity and the likelihood ratios between the probability measures of interest. Then, we make the link with ruin theory.

2.3.1. *The counting process N, and the intensity processes.* The counting process N takes values in $\mathbb{N}$ and its jump sizes are all equal to 1. It is defined on a measurable space $(\Omega, \mathcal{F}_\infty)$, equipped with a given filtration $\mathbb{F} = (\mathcal{F}_t)$ satisfying the usual conditions and a family of equivalent probability measures $\mathbb{P}^\theta$. The arrival epochs $(T_i)_{i\geq 1}$ are $\mathbb{F}$-stopping times. The reference probability measure is the nominal probability measure $\mathbb{P}$ when the change never occurs, i.e. $\theta = \infty$. The probability measure $\mathbb{P}^0$, associated with immediate change, i.e. $\theta = 0$, is also of great use. In the following, we adopt the notation of Kyprianou [18], where $\mathbb{P}^0$ is known as the tilded-probability associated with $\mathbb{P}$ and is denoted $\widetilde{\mathbb{P}}$. Henceforth, the tilded notation is referring to the quantities considered under $\widetilde{\mathbb{P}}$.

(i) Intensity process. As usual, the $\mathbb{F}$-adapted counting process N is characterized by its intensity process $\check{\lambda}_t$, varying with the reference probability measure $\mathbb{P}$ or $\mathbb{P}^\theta$. In particular, when N is a $\mathbb{P}$-Poisson process with time-varying independent increments, the intensity is a deterministic function λ_t giving the parameter of the Poisson distribution of N_t, with cumulative intensity parameter $\Lambda_t = \int_0^t \lambda_s ds$. Then, the Laplace transform of N_t is $\mathbb{E}[\exp(\alpha N_t)] = \exp(\Lambda_t(e^\alpha - 1))$, for $\alpha > 0$. Moreover, N is a strong Markov process such that for any finite stopping time S, the process $(N_{S+t} - N_S)$ is a Poisson process independent of $(N_t)_{t\leq S}$, and more generally of $\mathcal{F}_S$.

(ii) Martingale characterization of the intensity. When working with a stochastic intensity, we use a martingale to characterize the intensity process and a Wald martingale instead of the Laplace transform. The intensity process (λ_t) is now assumed to be (strictly) positive and adapted to the filtration $(\mathcal{F}_t)$, (predictable) with (strictly) increasing cumulative intensity process Λ_t. In other words, N is a so-called Cox-process or doubly stochastic counting process [14].

a) The process Λ_t is said to be the $\mathbb{P}$-cumulative intensity of the counting process N if the compensated process $M_t = N_t - \Lambda_t$ is a $\mathbb{P}$-martingale (if Λ_t is $\mathbb{P}$-integrable for any $t \geq 0$ and a local martingale if Λ_t is only finite a.s.). Sometimes, N is called a $(\mathbb{P}, \Lambda)$-counting process. So, by definition of $\widetilde{\mathbb{P}}$, N is a $(\widetilde{\mathbb{P}}, \tilde{\Lambda})$-counting process, such that $\tilde{\Lambda}_t = \rho\Lambda_t$. Then, the compensated process $\widetilde{M}_t = N_t - \tilde{\Lambda}_t$ is a $\widetilde{\mathbb{P}}$-(local) martingale.

b) The Esscher transform, generalizing the Laplace transform, is associated with the Wald (local) martingale $\mathcal{E}_t = \exp\big(\alpha N_t - \Lambda_t(e^\alpha - 1)\big)$. A more convenient form for our purpose is based on the exponential martingale denoted $\mathcal{E}^\eta_t$ and defined for $\eta > 0$ as

$$\mathcal{E}^\eta_t = \exp(\log(\eta)N_t - (\eta - 1)\Lambda_t) = \eta^{U^\eta_t}, \text{ where } U^\eta_t = N_t - \beta(\eta)\Lambda_t, \tag{5}$$

with $\beta(\eta) = (\eta-1)/\log(\eta) = \int_0^1 \eta^x dx$, $\eta \neq 1$, and $\beta(1) = 1$.
The differential form of (5) can easily be obtained using pathwise differential arguments,

$$d\mathcal{E}_t^\eta = (\eta-1)\mathcal{E}_{t-}^\eta(dN_t - d\Lambda_t) = (\eta-1)\mathcal{E}_{t-}^\eta dM_t. \tag{6}$$

Thus, any $(\mathbb{P},\Lambda)$-counting process N is characterized by the $\mathbb{P}$-local martingales $\mathcal{E}^\eta$, with $\eta > 0$, which are true martingales if Λ_t is bounded, or deterministic. In the general case (as in the continuous case), only sufficient conditions exist [19].

3. Log(ρ)-Sequential Probability Ratio and Surplus or Equity processes.

In this section, we introduce the reference processes in our study, and propose different points of view in their analysis, in particular via the connection with insurance theory.

3.1. Log_ρ*-sequential probability ratio process* U^ρ. We come back to the sequential detection problem in a proportional change between two (stochastic) intensity processes λ_t and $\rho\lambda_t$ using sequential probability ratio (SPR). As in the classical sequential probability ratio test , the main tool is the $\log_\rho$-SPR process.

3.1.1. *The* $\log_\rho$*-sequential probability ratios* $U_t^\rho = N_t - \beta(\rho)\Lambda_t$ *and* $X_t^\rho = -N_t + \beta(\rho)\Lambda_t$. We have seen in (5) that the process $(\rho^{U_t^\rho})$ is a local martingale with expectation equal to 1 in the deterministic case or when Λ_T is bounded. Under such an assumption, it is easy to define a probability measure $\widetilde{\mathbb{P}}$ on $\mathcal{F}_T$ by

$$d\widetilde{\mathbb{P}}/d\mathbb{P} = \mathcal{E}_T^\rho = \rho^{U_T^\rho}. \tag{7}$$

(i) In the inhomogeneous Poisson case, it is well-known that the $\widetilde{\mathbb{P}}$-intensity of N_t is $\rho\lambda_t$. To extend this property to the general case, we introduce the process $\widetilde{U}_t^\eta = N_t - \beta(\eta)\rho\Lambda_t$, and study the process $\tilde{\mathcal{E}}_t^\eta = \eta^{\widetilde{U}_t^\eta}$ under the probability measure $\widetilde{\mathbb{P}}$, or equivalently the process $\mathcal{E}_t^\rho\tilde{\mathcal{E}}_t^\eta = \exp(Z_t)$ under the probability measure $\mathbb{P}$. Since

$$\begin{aligned} Z_t &= \log(\rho)U_t^\rho + \log(\eta)\widetilde{U}_t^\eta, \\ &= \log(\rho\eta)N_t - \Lambda_t[(\rho-1)+(\eta-1)\rho)] = \log(\rho\eta)N_t - \Lambda_t(\rho\eta-1), \end{aligned}$$

the process $\exp(Z_t)$ is the $\mathbb{P}$-local martingale $\mathcal{E}_t^{\rho\eta}$, and $\tilde{\mathcal{E}}_t^\eta$ is a $\widetilde{\mathbb{P}}$-local martingale.
(ii) Similarly, $\mathbb{P}$ may be recovered from $\widetilde{\mathbb{P}}$ with the likelihood ratio $(\mathcal{E}_T^\rho)^{-1} = \rho^{-U_T^\rho}$. The remarkable property is that the $\widetilde{\mathbb{P}}$-local martingale $(\mathcal{E}_t^\rho)^{-1}$ is the $\widetilde{\mathbb{P}}$-local martingale $\tilde{\mathcal{E}}_t^{1/\rho}$,

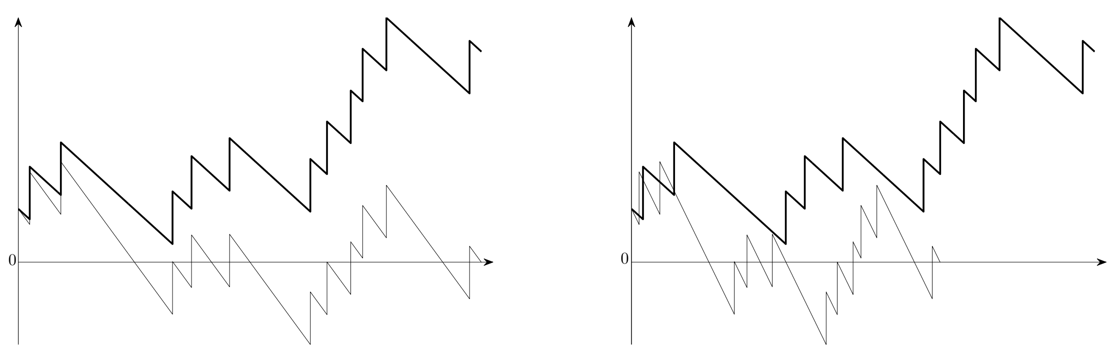
$$(\mathcal{E}_t^\rho)^{-1} = (1/\rho)^{(N_t-\beta(\rho)\Lambda_t)} = (1/\rho)^{(N_t-\beta(1/\rho)\rho\Lambda_t)} = \tilde{\mathcal{E}}_t^{\tilde{\rho}}, \tag{8}$$

with $\tilde{\rho} = 1/\rho$.
For convenience, we denote $X_t^\rho = -U_t^\rho$ the *dual* process of the log-likelihood ratio. Note that the process X^ρ has negative jumps and belongs to the large family of spectrally negative Lévy processes (Bertoin (1998) [9], Kyprianou (2006) [17]). In the sequel, we shall frequently drop out the subscripts ρ or β from the notation when there is no confusion.

3.1.2. *Time rescaling into a Poisson process.* There are two ways to change the intensity of the counting process N: by changing the probability measure as above, or by changing the time scale. The first one acts on the distribution of the counting process whereas the second one acts on its sample path. These two points of view yield to interesting interpretations of the same phenomenon.
(i) Time rescaling. We make the additional assumption that *the (strictly) increasing process* Λ_t

converges toward $+\infty$ *as* t *tends to* $+\infty$, *(*$\mathbb{P}$ *and* $\widetilde{\mathbb{P}}$*-a.s.)*. Then, the range of the inverse process Λ_t^{-1} is the interval $[0,\infty]$. For any $t \geq 0$, the r.v. Λ_t^{-1} is a stopping time with respect to the filtration $\mathbb{F}$. Denote by $\widehat{\mathbb{F}}$ the filtration $(\mathcal{F}_{\Lambda_t^{-1}})$. Hence it follows that, under the probability $\mathbb{P}$, the time rescaled and compensated counting process $\hat{N}_t - t = N_{\Lambda_t^{-1}} - t$ is a $\widehat{\mathbb{F}}$-local martingale. Therefore, $\hat{N}$ is a $(\mathbb{P}, \widehat{\mathbb{F}})$-Poisson process with intensity 1. Consequently, the interarrival times $(\widehat{T}_{j+1} - \widehat{T}_j)$ are independent with $\mathbb{P}$-exponential distribution with parameter 1. Finally, note that the counting process N may be obtained by the inverse time rescaling from the process $\hat{N}$. In particular, $N_t/\Lambda_t = \hat{N}_{\Lambda_t}/\Lambda_t$ goes to 1, $\mathbb{P}$-a.s. and to ρ, $\widetilde{\mathbb{P}}$-a.s. when t goes to $+\infty$.
(ii) Asymptotic behavior. Since $\hat{N}_t/t$ goes to 1, when t tends to ∞, $\mathbb{P}$-a.s, $\hat{U}^\rho$ drifts to $-\infty$, $\mathbb{P}$-a.s if $\beta > 1$, and $\hat{U}^\rho$ drifts to $+\infty$ *a.s* if $\beta < 1$. Then, $\rho^{U_t^\rho} = \rho^{\hat{U}^\rho_{\Lambda_t}}$ is a (local) martingale going to 0 at infinity, in any cases, but if $\beta(\rho) > 1$ (equivalently $\rho > 1$) its jumps are positive and negative when $\beta < 1$. The same property holds true after time rescaling by Λ_t/λ.

3.1.3. *Pathwise comparison.* Using these ideas, we obtain a way to simulate in the deterministic case, on the same probability space from the $(\mathbb{P}, \lambda)$-Poisson process N, the process $(U_t^\rho = N_t - \beta(\rho)\lambda t)$, and another process $\widetilde{\mathbb{P}}$-distributed as (U_t^ρ). This is illustrated in Figure 1.



(a) Processes U_t^ρ and $U_t^{1/\rho}$

(b) Processes U_t^ρ and $U_{\rho t}^{1/\rho}$

Figure 1. Sample paths, for $\rho > 1$, of the cusum processes U_t^ρ (left, thick line) and $U_t^{1/\rho}$ (left, thick line) as well as the processes U_t^ρ (right, thin line) and $U_{\rho t}^{1/\rho}$ (right, thick line) when λ is time-homogeneous (set equal to 1).

Proposition 1 (Pathwise equivalence)**.** *Assume* Λ *deterministic. Put* $\tilde{\rho} = 1/\rho$. *Recall that the* $(\widetilde{\mathbb{P}}, \tilde{\lambda})$*-process* N *has the same distribution as the* $(\mathbb{P}, \lambda)$*-process* $(N_{\rho t})$.
(i) *The* $(\widetilde{\mathbb{P}}, \tilde{\lambda})$*-process* $(U_t^\rho = N_t - \beta(\rho)\Lambda_t)$ *has the same distribution as the* $(\mathbb{P}, \lambda)$*-process* $(U_{\rho t}^{\tilde{\rho}})$.
(ii) *Observe that* $U^\rho < U^{\tilde{\rho}}$ *if* $\beta(\rho) > 1$, *and* $U^\rho \geq U^{\tilde{\rho}}$ *if* $\beta(\rho) < 1$. *So, the process* $(\mathbb{P}, U_t^\rho)$ *is stochastically dominated by the process* $(\widetilde{\mathbb{P}}, U_{\tilde{\rho}t}^\rho)$ *when* $\rho > 1$. *That is, for any increasing function* G *on the space of the paths,* $\mathbb{E}(G(U^\rho)) \leq \widetilde{\mathbb{E}}(G(U_{\tilde{\rho}.}^\rho))$.

Proof. (i) By the scaling property, the $(\widetilde{\mathbb{P}}, \rho\lambda)$-Poisson process N has the same distribution as the rescaled $(\mathbb{P}, \lambda)$-Poisson process $N_{\rho t}$. So, the $\widetilde{\mathbb{P}}$-distribution of $(U_t^\rho = N_t - \beta(\rho)\lambda t)$ is the same as the $\mathbb{P}$-distribution of $N_{\rho t} - \beta(\rho)\lambda t = N_{\rho t} - (\beta(\rho)/\rho)\rho\lambda t = U_{\rho t}^{\tilde{\rho}}$, $(\beta(\rho) = \rho\beta(\tilde{\rho}))$.
(ii) Assume for example $\rho > 1$. The two processes U^ρ and $U^{\tilde{\rho}}$ are ordered, since $U_t^\rho \leq U_t^{\tilde{\rho}}$, for

all $t \geq 0$. Since the process $(U^{\tilde{\rho}}_{\rho t})$ is distributed as (U^{ρ}_t) under $\widetilde{\mathbb{P}}$, then $(U^{\tilde{\rho}}_t)$ is distributed as $(U^{\rho}_{\tilde{\rho}t})$ under $\widetilde{\mathbb{P}}$. So, this pathwise inequality implies the stochastic dominance of $(\mathbb{P}, U^{\rho}_t)$ by $(\widetilde{\mathbb{P}}, U^{\rho}_{\tilde{\rho}t})$. In particular, for any increasing function G on the space of the paths, $\mathbb{E}(G(U^{\rho})) \leq \widetilde{\mathbb{E}}(G(U^{\rho}_{\tilde{\rho}.}))$. $\square$

3.2. *Surplus and equity processes of an insurance company.*
Here, we make an interesting connection with ruin theory about which an abundant literature exists, see Rolski et al. [31] among many others.

3.2.1. *Surplus and equity processes.* (i) a) SURPLUS PROCESS. Let us consider an insurance company, receiving premiums at rate $(\rho-1)\lambda_t$ (with $\rho > 1$), when the number of claims arriving up to time t is the counting process N_t, with intensity λ_t and average cost by claim $\log\rho$. The surplus process X is written using the average cost by claim as numéraire with initial reserve z, as $X_t(z) = z + X_t$. Then, the cash value of the surplus process is $(\log\rho)\, X_t(z) = (\log\rho)(z - N_t) + (\rho-1)\Lambda_t$, and

$$X_t(z) = z + X_t = z - N_t + \beta(\rho)\Lambda_t, \quad \text{with } \rho > 1. \tag{9}$$

b) EQUITY PROCESS. On the other hand, there is the so-called dual risk model. For such a process, premia can be regarded as costs and claims as profits, and the *surplus* can be interpreted as a capital of an economic activity where the gains come from suddenly as in research and development (Avanzi et al. [4] (2007), Afonso et al. [1] (2013)). So, the viability condition becomes $\mathbb{E}(U_1) > 0$ or $\beta(\rho) < 0$. As above, using the average gain by success as numéraire, the equity process with initial capital z is the process $z + U^{\rho}_t$, associated with a cash value $(\log\rho)\, U_t(z)$,

$$U_t(z) = z + U_t = z + N_t - \beta(\rho)\Lambda_t, \quad \text{with } \rho < 1. \tag{10}$$

(ii) RUIN OR PERFORMANCE PROBLEM. A classical problem in insurance theory is the so-called *ruin problem*. That is the computation, under *the security loading condition* $\beta > 1$ of the ruin probability in infinite time given an initial capital $z > 0$, that is $\mathbb{P}(\exists\, t \text{ s.t. } X_t \leq -z)$. The problem may be also formulated in terms of the dual process $U = -X$, as $\mathbb{P}(\exists\, t \text{ s.t. } U_t \geq z) = \mathbb{P}(\bar{U}_\infty \geq z)$, where $\bar{U}_\infty = \sup_t U_t$. We have seen that $\bar{U}_\infty$ is finite a.s. Its distribution function $\bar{u}(m) = \mathbb{P}(\bar{U}_\infty \leq m)$ is well known for a long time [2]. Its analytic closed form is recalled in Theorem 5. In the equity problem $(\beta < 1)$, the ruin occurs when the minimum of the equity process attains 0. But, the main problem concerns the performance level of U, that is the probability to attain some capital level m (before the ruin.). Obviously, the time to ruin, that is the first time where X_t goes below the level $-z$ is also of particular interest.

(iii) CROSSING HIGH OR LOW BARRIERS AND RUNNING EXTREMA. Since we are concerned with different processes, we give the definitions for a general right continuous (continu à droite), left limited (limité à gauche) (càdlàg) finite variation process Z_t with finite number of positive (negative) jumps.
The general definition is the following, where as usual the infimum of the empty set is $+\infty$,

$$\tau^Z_m = \inf\{t : Z_t \geq m\}, \quad \text{and} \quad \sigma^Z_b = \inf\{t : Z_t \leq b\}. \tag{11}$$

With this definition, if $Z_0 \geq m$, $\tau^Z_m = 0, a.s.$ and if $Z_0 \leq b$, $\sigma^Z_b = 0, a.s.$. The family $(\tau^Z_m)_m$ is non-decreasing, and the family $(\sigma^Z_b)_b$ is non-increasing. Their inverses are easily expressed in terms of the running supremum $\bar{Z}_t = \sup_{s\leq t} Z_s$ or minimum $\underline{Z}_t = \inf_{s\leq t} Z_s$, since $(\tau^Z_m \geq t \Leftrightarrow \bar{Z}_t \leq m)$ and $(\sigma^Z_b \geq t \Leftrightarrow \underline{Z}_t \geq b)$. In particular, $\{\tau^Z_m = +\infty\} = \{\bar{Z}_\infty \leq m\}$ and $\{\sigma^Z_b = +\infty\} = \{\underline{Z}_\infty \geq b\}$.

(iv) MARTINGALE PROPERTY AND SCALE FUNCTIONS. Since the process $U_t = U_0 + N_t - \beta\Lambda_t$ has positive jumps and a decreasing drift ($\beta > 0$), the barrier $b = 0 < U_0$ is only crossed continuously at time σ_0^U by the process $U_.(x)$ and on $\{\sigma_0^U < +\infty\}$, we have $U_{\sigma_0^U} = 0$. Put $\mathbb{P}_x(A) = \mathbb{P}(A|U_0 = x)$. Then,

Proposition 2. *Assume $\beta > 1$.*

(i) *The function $\theta_m(x) = \mathbb{P}_x(\sigma_0^U < \tau_m^U)$ is equal to $\theta_m(x) = \bar{u}(m-x)/\bar{u}(m)$ on $(0,m)$, and $\theta_m(U_t)$ is a $\mathbb{P}_x$-martingale on $[0, \sigma_0^U \wedge \tau_m^U)$.*

(ii) *Put $\tilde{u}(x) = \rho^x\,\bar{u}(x)$. Then, $\tilde{\theta}_m(x) = \widetilde{\mathbb{P}}_x(\sigma_0^U < \tau_m^U) = \tilde{u}(m-x)/\tilde{u}(m)$ and $\tilde{\theta}_m(U_t)$ is a $\widetilde{\mathbb{P}}_x$-martingale on $[0, \sigma_0^U \wedge \tau_m^U)$.*

Proof. When $\beta > 1$, it is well known that $\mathbb{P}_x(\sigma_0^U < \infty) = 1$, $\mathbb{P}_x(\bar{U}_\infty < \infty) = 1$ and $\mathbb{P}_x(\bar{U}_\infty \leq m) = \bar{u}(m-x) = \mathbb{P}_x(\tau_m^U = \infty)$. Then

$$\mathbb{P}_x(\tau_m^U = \infty) = \mathbb{P}_x(\{\tau_m^U > \sigma_0^U\} \cap \{\tau_m^U = \infty\}) = \mathbb{P}_x(\tau_m^U > \sigma_0^U)\,\mathbb{P}_{x=0}(\tau_m^U = \infty).$$

Then, for $x \in (0,m)$, $\mathbb{P}_x(\tau_m^U > \sigma_0^U) = \bar{u}(m-x)/\bar{u}(m) := \theta_m(x)$. By the Markov property, $\theta_m(U_t) = \mathbb{P}_x(\tau_m^U > \sigma_0^U|\mathcal{F}_t)$ on $[0, \sigma_0^U \wedge \tau_m^U)$ is a $\mathbb{P}_x$-martingale.
Therefore, under the probability $\widetilde{\mathbb{P}}_x$ with density ρ^{U_t}, the process $\rho^{-U_t}\theta_m(U_t)$ is a martingale on $[0, \sigma_0^U \wedge \tau_m^U)$ with terminal value $\mathbf{1}_{\{\tau_m^U > \sigma_0^U\}}$. So, it is natural to put $\tilde{u}(x) = \rho^x \bar{u}(x)$, so that $\widetilde{\mathbb{P}}_x(\tau_m^U > \sigma_0^U) = \tilde{u}(m-x)/\tilde{u}(m)$. □

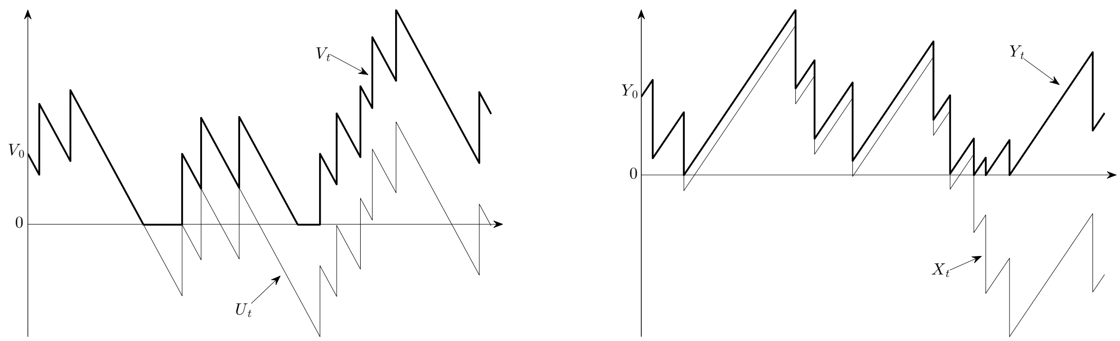
In Section 5, we will show that the functions $\bar{u}$ and $\tilde{u}$ belong to the family of scale functions (Bertoin [9]), that are proportional to the inverse of the Laplace exponent $\psi^\rho(\alpha) = \log(\mathbb{E}(\exp(-\alpha U_1^\rho)) = \lambda(\alpha\beta(\rho) + e^{-\alpha} - 1) = \lambda\alpha(\beta(\rho) - \beta(e^{-\alpha}))$ and $\tilde{\psi}^\rho(\alpha) = \log(\widetilde{\mathbb{E}}(\exp(-\alpha U_1^\rho)) = \lambda\alpha(\beta(\rho) - \rho\beta(e^{-\alpha})) = \rho\psi^{\tilde{\rho}}(\alpha)$.

4. Reflected counting process with drift.

We now introduce the key notions for the detection problem and also make the connection with the similar notions in insurance. The main processes of interest is the reflection of U and X respectively on their minimum and maximum. As it is emphasized by Pistorius (2004) [28], in applied probability the reflected processes also occur in the study of water level in a dam, in queuing theory (Asmussen (2003) [2], Prabhu (1998) [30]) and more recently in finance, for the study of Russian options as in Peskir and Shyriaev (2002) [26].

4.1. *Cusum process and surplus or equity process with dividends.* We start with the definition of the cusum process as a running maximum of the log-likelihood process. The definition depends on the sense of variation of ρ^z or equivalently of the sign of $\rho - 1$. The corresponding concepts in insurance are the surplus or the equity processes with dividends.

4.1.1. *Cusum process and reflexion.* As usual in estimation theory, the process of interest is the maximum likelihood process; in detection theory, since the parameter is a date, the maximum must be taken over the dates and the process of interest is $\sup_{\theta \leq t} \rho^{U_t - U_\theta}$.
(i) When $\rho > 1$, the maximum likelihood process is given by ρ^{V_t}, where $V_t = U_t + \sup_{s \leq t}(-U_s) = \bar{X}_t - X_t$, which is the process U_t reflected at 0, or the process X reflected at its maximum. When $\rho < 1$, the function ρ^z is decreasing and the maximum likelihood process is the process $(1/\rho)^{Y_t}$, where $Y_t = \sup_{\theta \leq t} U_\theta - U_t = \bar{U}_t - U_t = X_t + \bar{U}_t$. The process Y_t is the process U reflected at its maximum, or the process X reflected at 0.
(ii) The process $V_t = \bar{X}_t - X_t$ is well-defined for any $\rho > 0$, but a cusum process only when $\rho > 1$. In the Poisson case, the process $V_t = \bar{X}_t - X_t$ may be viewed as the Cramer-Lundberg



(a) Case $\rho > 1$

(b) Case $\rho < 1$

FIGURE 2. Sample paths of the cusum processes V (left, thick line) and Y (right, thick line) as well as the associated processes U and X (thin lines) when the intensity λ is constant.

process with negative jumps X_t reflected at its continuous maximum $\bar{X}_t$, see Bertoin [9] and Kyprianou [17] for instance. In particular, when λ is constant, the *symmetry principle* allows us to identify the distribution of the two random variables $V_t = (\bar{X}_t - X_t)$ (with t being fixed) and $-\inf_{s\leq t} X_t = -\underline{X}_t = \bar{U}_t$. In particular, $\mathbb{P}(\bar{U}_t \leq m) = \mathbb{P}(\tau_m^U \geq t) = \mathbb{P}(V_t \leq m)$, and $\mathbb{E}(e^{-q\tau_m^U}) = \mathbb{E}\int_0^\infty q\, e^{-qt}\mathbf{1}_{\{V_t\leq m\}}dt$. This property may be extended at any independent exponential time e_q, with mean $1/q$, yielding to useful developments, in particular the famous Wiener-Hopf decomposition linking the maximum and minimum up to e_q.

(iii) The process $Y_t = \bar{U}_t - U_t$ is well-defined for any $\rho > 0$, but a cusum process only when $\rho < 1$. The process $Y_t = \bar{U}_t - U_t$ is the process U reflected at its maximum. Since the process U is decreasing between two jumps of N, the process $\bar{U}_t$ is not continuous, and the epochs when a new supremum of $\bar{U}$ is reached are times arrival of N, for which $Y_t = 0$. As before, in the Poisson case, the variables $Y_t = \bar{U}_t - U_t = \sup_{s\leq t}(U_s - U_t) = \sup_{s\leq t}(X_t - X_s)$ and $\bar{X}_t$ have the same distribution.

(iv) For a generic initial condition Z_0, the associated cusum processes become the processes $V_t(Z_0)$ if $\rho > 1$, and $Y_t(Z_0)$ if $\rho < 1$ given by

$$V_t(Z_0) = Z_0 + U_t + \sup\{Z_0, \bar{X}_t\} = U_t(Z_0) + (\bar{X}_t - Z_0)^+ = V_t + (Z_0 - \bar{X}_t)^+, \quad (12)$$

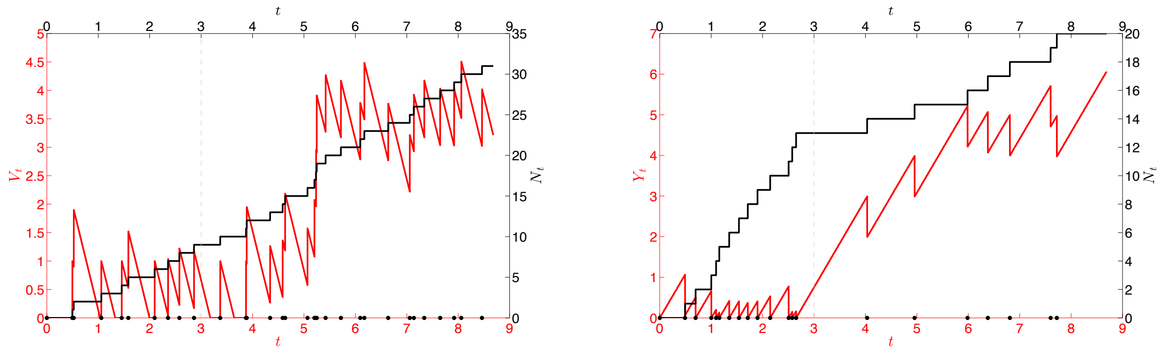
$$Y_t(Z_0) = Z_0 + X_t + \sup\{Z_0, \bar{U}_t\} = X_t(Z_0) + (\bar{U}_t - Z_0)^+ = Y_t + (Z_0 - \bar{U}_t)^+, \quad (13)$$

$$\bar{X}_t^{ad}(Z_0) = (\bar{X}_t - Z_0)^+, \quad \bar{U}_t^{ad} = (\bar{U}_t - Z_0)^+. \quad (14)$$

With this notation, the increasing processes $\bar{X}_t^{ad}(Z_0)$ and $\bar{U}_t^{ad}(Z_0)$ are still starting from 0. In the Poisson case, the increasing processes $\bar{X}_t^{ad}$ and $\bar{U}_t^{ad}$ are additive functionals of the processes V and Y.

(v) SAMPLE PATHS. We see in Figure 2 that the process V behaves as $U(V_0)$ above 0, jumping with size 1 at times $(T_i)_{i\geq 1}$. In between jumps, V decreases at rate $\beta\lambda$ which is cut off when V reaches 0, i.e. V is reflected at 0. The process Y behaves as $X(Y_0)$ above 0, jumping at times $(T_i)_{i\geq 1}$. In between jumps, $Y(x)$ increases at rate $\beta\lambda_.$, but its jumps are cut-off (reflection) when X_t is below 0 in such a way than $Y_t = 0$.

Figure 3(a) shows a simulated path of the Poisson process N_t with intensity 3 up to time $\theta = 3$ and 4.5 afterwards ($\rho = 1.5$), as well as the associated sample path of $V_t(0)$. In the same manner, we depict in Figure 3(b) a simulation of a decrease in the intensity from $\lambda = 3$ to 1.5 ($\rho = 0.5$): we plot the associated cusum process $Y_t(0)$ together with the counting process N_t.



(a) Case $\rho > 1$

(b) Case $\rho < 1$

FIGURE 3. Simulated cusum processes V and Y with a change-point at time $\theta = 3$ and a constant intensity 3. The post change parameter ρ is set equal to 1.5 and 0.5 respectively. The left axis of ordinates is associated with the cusum process and the right axis is related to non-decreasing process N. The dots in the abscissa indicate the jump epochs T_i.

4.1.2. *Insurance point of view.* (i) SURPLUS PROCESS WITH DIVIDENDS. We have seen that the surplus process X of an insurance company, using the average cost by claim $|\log \rho|$ as numéraire, and initial reserve z, is $X_t(z) = z + X_t = z - N_t + \beta(\rho)\Lambda_t$.
We now consider that the insurance company pays out dividends (De Finetti (1957) [11]) at the same rate $|\rho - 1|\lambda_t$, when the cash reserves come above a target level $m|\log \rho|$. Then, the dynamics of the surplus process $\hat{X}_t(z,m) = \hat{X}_t^{(m)}$ is driven by the following ordinary differential equation (ODE for short), for $0 < z < m$, with initial condition $\hat{X}_0(z,m) = z < m$,

$$d\hat{X}_t^{(m)} = -dN_t + \beta \mathbf{1}_{(0,m)}(\hat{X}_t^{(m)})d\Lambda_t, \quad z < m. \tag{15}$$

The main point established in the following proposition is that the surplus process with dividends $(\hat{X}_t^{(m)})$ is a linear transform of the reflected process $(V_t(z))$ on the interval $0 < z < m$. Mathematically speaking, this property expresses that the continuous increasing process $\bar{X}_t$ (whose support is the set $\{t; \bar{X}_t - X_t = V_t = 0\}$) is absolutely continuous with respect to the Lebesgue measure.

Theorem 3. (i) *The reflected process of* $U_t(V_0) = V_0 + N_t - \beta\Lambda_t$ *at* 0 *is the process* $V_t = U_t(0) + \sup(V_0, \bar{X}_t) = U_t(V_0) + \bar{X}_t^{ad}(V_0)$ *where* $\bar{X}_t^{ad}$ *is a continuous increasing process.*
(ii) V *is the unique solution of the ODE driven by the counting process* N,

$$V_t = V_0 + N_t - \beta \int_0^t \mathbf{1}_{(0,\infty)}(V_s)d\Lambda_s, \text{ that implies } \bar{X}_t^{ad} = \int_0^t \beta \mathbf{1}_{\{V_s=0\}}d\Lambda_s. \tag{16}$$

In other words $\mu - V_t = \hat{X}_t^{(\mu)}(\mu - V_0)$ *is a surplus process in ruin theory.*
(iii) *When* Λ *is deterministic, the process* (V_t) *is a homogeneous strong Markov process.*

Proof. a) The result is a consequence of a classical result based on differential calculus. That is, the continuous part of a right continuous finite variation process V does not charge the set $\{V_. = 0\}$. Since the continuous part of V_t is $-\beta\Lambda_t + \bar{X}_t^{ad}$,

$$d\bar{X}_t^{ad} = \beta \mathbf{1}_{\{V_t=0\}} d\Lambda_t.$$

So, we have shown that V is solution of the ODE (16) together with the representation of the supremum $\bar{X}^{ad}$ in differential form.

b) If V and V' are two non-negative solutions starting from x, the difference is a continuous differentiable process,

$$d(V_t - V_t')^+ = \mathbf{1}_{\{V_t > V_t'\}}(\mathbf{1}_{\{V_t>0\}} - \mathbf{1}_{\{V_t'>0\}})(-\beta)d\Lambda_t = \mathbf{1}_{\{V_t>0, V_t'=0\}}\beta d\Lambda_t.$$

The process $(V_t - V_t')^+$ starting from 0 is non-negative and non-increasing; so is the null process. Interverting the role of V and V', we obtain that $V_t \equiv V_t'$ and so the uniqueness.

c) When Λ_t is deterministic, the strong Markov property can be easily deduced from the uniqueness of the equation. □

(ii) Equity process with dividends. For the equity process, we still have an upper constant dividend barrier m strategy, and the ruin level at 0. An immediate amount of surplus in excess of m is paid in the form of a dividend, see for e.g. Afonso et al. [1], Avanzi et al. [4]. Hence, the process restarts at level m if this is overtaken by a claim. So, the process is reflected at the level m and the $m-$ equity process is reflected at 0 with ruin when it hits m. There is a significant difference in the differential representation of processes V and Y due to the non-continuity of the process $\bar{U}$. Nevertheless, we can give a differential representation as in (16), for the process Y with a discontinuous local time $\bar{U}^{ad}$. The proof relies on elementary finite variation differential calculus, recalled in the next subsection.

Theorem 4. *The reflected process of $X_t(Y_0) = Y_0 - N_t + \beta\Lambda_t = Y_0 - U_t$ at 0 is the non-negative process $Y_t(Y_0) = X_t(0) + \sup(Y_0, \bar{U}_t) = X_t(Y_0) + \bar{U}_t^{ad}(Y_0)$. The jumps of the process Y are still at some epochs of the jumps of N, with, in addition, the property that $Y_t = 0$.*

(i) *The process $\bar{U}_t^{ad} = (\bar{U}_t - Y_0)^+$ is an increasing pure jump process, with differential*

$$\begin{cases} d\bar{U}_t^{ad} &= (U_t - \bar{U}_{t-})^+ dN_t = \mathbf{1}_{\{Y_t=0\}}(1 - Y_{t-})dN_t, \quad \bar{U}_0^{ad} = 0, \\ dY_t &= -j(Y_t)dN_t + \beta d\Lambda_t, \quad j(y) = y \wedge 1. \end{cases} \tag{17}$$

(ii) *Y is the unique solution of the ODE (16), driven by the counting process N. In the Poisson case, Y is a strong Markov process.*

Proof. (i) The relation $d\bar{U}_t^{ad} = (U_t - \bar{U}_{t-})^+ dN_t$ expresses that $\bar{U}_t^{ad}$ is an increasing pure jumps process, with jumps size $U_t - \bar{U}_{t-} = 1 - Y_{t-}$ since $Y_t = 0$. Since the process U jumps by 1 as N does, by the definition of $Y = \bar{U}_t^{ad} - U$, we see that

$$dY_t = -(Y_{t-}\mathbf{1}_{\{Y_t=0\}} + \mathbf{1}_{\{Y_t>0\}})dN_t + \beta d\Lambda_t = -j(Y_{t-})dN_t + \beta d\Lambda_t,$$

where $j(x) = x \wedge 1$.

The last equality $(Y_{t-}\mathbf{1}_{\{Y_t=0\}} + \mathbf{1}_{\{Y_t>0\}}) = j(Y_{t-})$ is true, since the only case where a jump of Y at time t is different of 1 is when $Y_t = 0$.

(ii) If Y and Y' are two non-negative solutions starting from x, the difference Φ_t is a pure jump process, starting from 0 such that $d\Phi_t = d(Y_t - Y_t') = (j(Y_{t-}) - j(Y_{t-}'))dN_t$. On the interval $[0, T_1)$, $\Phi_t = 0$, and $j(Y_t) - j(Y_t')) = 0$. Then, at T_1 the jump of Φ_t is 0, and $\Phi_{T_1} = 0$. Using the same argument between the successive dates of jumps of N, we see that the process Φ_t is still equal to 0. □

4.1.3. *Cusum performances of reflected processes.* (i) Cusum performances. The cusum rule is defined as the first time $\tau_m^{\rho,\mathrm{cus}}$ where the cusum process V or Y exceeds a given level $m>0$, and the cusum performance is measured via the cusum criterion $\{\widetilde{\mathbb{E}}(N_{\tau_m^{\rho,\mathrm{cus}}}-N_\theta\big|\mathcal{F}_\theta),\ \theta<\tau_m^{\rho,\mathrm{cus}}\}$, and the false alarm constraint is expressed with the help of $\mathbb{E}(N_{\tau_m^{\rho,\mathrm{cus}}})$. Using a change of time argument, we are concerned in the Poisson case with the functions $\widetilde{C}_m^{\mathrm{cus}}(x)=\widetilde{\mathbb{E}}_x(N_{\tau_m^{\rho,\mathrm{cus}}})$ and $C_m^{\mathrm{cus}}(x)=\mathbb{E}_x(N_{\tau_m^{\rho,\mathrm{cus}}})$, for $0\le x<m$, since outside these functions are equal to 0.
The *pathwise equivalence* (Proposition 1) allows us to work only under the probability $\mathbb{P}$ (or $\mathbb{P}_x$) with the cusum processes with parameters ρ and $\tilde\rho=1/\rho$, and their stopping times $\tau_m^{\rho,\mathrm{cus}}$ and $\tau_m^{\tilde\rho,\mathrm{cus}}$, since $\widetilde{\mathbb{E}}_x(N_{\tau_m^{\rho,\mathrm{cus}}})=\mathbb{E}_x(N_{\tau_m^{\tilde\rho,\mathrm{cus}}})$. Moreover, the distinction between cusum processes and reflected processes is not necessary before proceeding to the optimality results.
(ii) Links with reflected processes. As described in Theorems 3 and 4, the reflected processes V and Y do not cross upper barrier m with the same regularity, since V increases only by jumps Y only continuously. So, these two processes behaves differently.
a) In what follows, for a given ρ, τ_m^V and τ_m^Y denotes the hitting times of m by the processes V_t^ρ and Y_t^ρ, while $\tilde\tau_m^V$ and $\tilde\tau_m^Y$ denotes the hitting time of m by the processes $V_t^{\tilde\rho}$ and $Y_t^{\tilde\rho}$, with $\tilde\rho=1/\rho$. The performance functions are denoted $h_m(x)=\mathbb{E}_x(N_{\tau_m^V})$, $g_m(x)=\mathbb{E}_x(N_{\tau_m^Y})$, and $\tilde h_m(x)=\mathbb{E}_x(N_{\tilde\tau_m^V})=\widetilde{\mathbb{E}}_x(N_{\tau_m^V})$, $\tilde g_m(x)=\widetilde{\mathbb{E}}_x(N_{\tilde\tau_m^Y})=\widetilde{\mathbb{E}}_x(N_{\tau_m^Y})$.
Although τ_m^U may be infinite when $\rho>1$, it has been shown that τ_m^V is finite $\mathbb{P}_x$-a.s. for any $m\ge1$. Similarly, and σ_0^X may be infinite when $\rho<1$, but it has been shown that σ_0^Y is finite $\mathbb{P}_x$-a.s. for any $x>0$. Different proofs are given depending on the context: in the ruin theory, the finiteness of the time to ruin with dividends can be found in Asmussen and Albrecher [3, Ch. VIII] for instance, whereas in the reliability theory more complete results is given in Zacks [36].
b) In the Poisson case, thanks to the Markov property of the reflected processes, the performance functions are associated with different martingales, defined on $[0,\tau_m^{\mathrm{cus}})$,

$$\begin{aligned}
&\text{Under } \mathbb{P}_x, && H_t^m=N_t+h_m(V_t)-h_m(V_0),\ G_t^m=N_t+g_m(Y_t)-g_m(Y_0). \\
&\text{Under } \widetilde{\mathbb{P}}_x, && \widetilde{H}_t^m=N_t+\tilde h_m(V_t)-\tilde h_m(V_0),\ \widetilde{G}_t^m=N_t+\tilde g_m(Y_t)-\tilde g_m(Y_0).
\end{aligned}\tag{18}$$

But the restriction to the Poisson case is not necessary, since the martingale property is stable by time rescaling. Moreover, below, we will use differential calculus to extend the martingale property at all $\mathbb{R}^+$.
(iii) Pathwise comparison. Thanks to identity $\beta(\rho)=1/\log(\rho)\int_0^1\rho^x\,dx$, the functions $\rho\mapsto\beta(\rho)$ and $\rho\mapsto -U_t^\rho=\beta(\rho)\Lambda_t-N_t$ are increasing for $\rho>1$, and decreasing for $\rho<1$. The same properties hold true for their variations across any interval $]s,t]$. Therefore, for any $s<t$ and $\tilde\rho\le\rho$, $U_t^\rho-U_s^\rho<U_t^{\tilde\rho}-U_s^{\tilde\rho}$, and $V_t^\rho(0)<V_t^{\tilde\rho}(0)$ when $Y_t^\rho(0)>Y_t^{\tilde\rho}(0)$. Since, $V_t^\rho(x)-V_t^\rho(0)=(x-\bar X_t^\rho)^+$, and $\rho\mapsto\bar X_t^\rho$ is increasing, the process $\rho\mapsto V_t^\rho(x)$ is decreasing in ρ, and the family of stopping times $(\tau_m^{\rho,V}(x))$ is increasing in $\rho>1$. The same inequalities applied at $Y_t(y)=\sup_{s\le t}(U_s-U_t)$ yield to reverse inequalities. So, we obtain the inequalities,

$$\begin{cases}\tilde h_m^\rho(x) &= h_m^{\tilde\rho}(x)\le h_m^\rho(x), \quad \text{when } \rho>1,\\ \tilde g_m^\rho(x) &= g_m^{\tilde\rho}(x)\ge g_m^\rho(x), \quad \text{when } \rho<1.\end{cases}\tag{19}$$

5. Differential calculus, Delayed equation, and Martingale.

The use of the finite variation calculus allows us to characterize the scale functions, and the performance functions from their martingale property. In the Poisson case, stochastic differential formula yields to show that these functions are solutions of a delayed differential equation (DDE in short) defined on the domain $(0,m)$. The latter is similar to the one satisfied by the scale functions $\bar{u}$ and $\tilde{u}$ defined in Proposition 2, but with drift and first order boundary condition at 0 induced by the reflection at 0. Such a condition is said to be a Neumann condition. An additional difficulty comes from discontinuities of the performance functions at m. Fortunately, we only need ordinary finite variation differential calculus.

5.1. *Finite variation calculus, delay equation and Pollaczek-Khintchine formula.* Recall the deterministic differential rule for a càdlàg finite variation function ϕ on $\mathbb{R}^+$, with finite number of jumps $\delta\phi(s) = \phi(s) - \phi(s-)$, and a.e. differentiable with left-hand derivative ϕ',

$$\phi(z) = \phi(0) + \int_0^z \phi'(u)du + \sum_{\alpha \le z} \phi(\alpha) - \phi(\alpha-). \tag{20}$$

In distribution theory, the function ϕ is differentiable, with distribution differential,

$$\phi(du) = \phi'(u)du + \sum_{\{\alpha, \delta\phi(\alpha)\neq 0\}} \delta\phi(\alpha)\, \delta_\alpha(du), \quad \delta\phi(x) = \phi(x) - \phi(x-).$$

Before proceeding to the introduction of the differential formula in stochastic calculus, we give an immediate application of the above differential rule to delayed equations.

5.1.1. *Finite variation calculus and delay differential equation.* An immediate application is the study of delayed differential equation by a fixed point method. Special attention is paid to the simplest form of the delayed equation (21), whose solutions are the scale functions, $\bar{u}(x)$ and $\tilde{u}(x)$ introduced in Proposition 2. As explained in the next subsection, the martingale property of these functions taken at $m - U^\rho_\cdot$ on the domain $(0,m)$ is enough to show that they satisfy the following DDE.

Theorem 5. *Let us consider the delayed equation on $[0,\infty)$, whose finite variation non-negative solution u, (null on $[-\infty,0)$), is continuous with only a jump at 0, $u(0) > 0$,*

$$\beta u'(x) = u(x) - u(x-1), \quad \beta > 0. \tag{21}$$

This delayed equation is stable under some simple transformations,
(i) a) *Assume $\beta = \beta(\rho)$. If u is solution of (21), then $\rho^x u(x)$ is solution of (21) with new coefficient $\tilde{\beta}(\rho) = \beta(\rho)/\rho = \beta(1/\rho) := \beta(\tilde{\rho})$.*

b) *If u is solution of (21), then $\hat{u}(x) = \int_0^x u(z)dz$, equal to 0 for $x \le 0$, is solution of the delayed differential equation with drift,*

$$\beta \hat{u}'(x) = \hat{u}(x) - \hat{u}(x-1) + \beta u(0) \quad \beta > 0, \quad u(0) = \hat{u}'(0). \tag{22}$$

(ii) *The derivative is a solution (null for $x < 0$) of the convolution equation,*

$$u'(x) = (1/\beta)\, \mathbf{1}_{[0,1)}(x) u(0) + (1/\beta) \int_0^1 u'(x-z)dz \quad a.e. \tag{23}$$

(iii) Pollaczek-Khintchine formula. *Assume $\beta > 1$. Let (U_i) be an i.i.d. sample of a uniformly distributed r.v. on $[0,1]$ with sum $S_n = \sum_{i=1}^n U_i$, , and ν an independent geometrical r.v. with distribution on $j \ge 0$ given by $\mathbb{P}(\nu = j) = (1 - 1/\beta)\beta^{-j}$. Then,*

$$u'(x) = u(0)/(\beta - 1)\mathbb{P}(S_\nu \in [x-1, x)) = u(0)(W(x) - W(x-1)). \tag{24}$$

where $W(x) = 1/(\beta-1)\mathbb{P}(S_\nu \le m)$, *and* $W(0) = 1/\beta$.
(iv) *When* $\beta > 1$, *the* $\bar{U}_\infty$*-cumulative distribution function* $\bar{u}(x) = \mathbb{P}(\bar{U}_\infty \le x)$ *is the solution of the delayed equation equal to* $\bar{u}(x) = \mathbb{P}(S_\nu \le x)$.

Proof. (i) a) The derivative of $\rho^x\, u(x) = u_\rho(x)$ is $u'_\rho(x) = \log(\rho)\rho^x\, u(x) + \rho^x\, u'(x)$. The relation $\beta(\rho)\log(\rho) = \rho - 1$, and some algebra yield to the equality

$$(\beta(\rho)/\rho)u'_\rho(x) = \beta(\tilde{\rho})u'_\rho(x) = u_\rho(x) - u_\rho(x-1).$$

b) The integral equation is based on formula (27) given that $\int_0^x u'(z)dz = u(x) - u(0)$. Then, the delayed equation of the primitive function $\hat{u}$ differs from the previous one by the addition of the constant $u(0)\beta$. The function $W(0)$ by definition satisfies $W(0)\beta = 1$.
(ii) The delayed function $u(x) - u(x-1)$ is continuous on $(0,\infty)$ outside of $x = 1$ and $x = 0$, with jumps of size $-u(0)$ and $u(0)$. By the differential formula (27) ,

$$u(x+1) - u(x) = u(0)\mathbf{1}_{[0,1)}(x) + \int_{(x-1)^+}^{x} u'(y)dy, = u(0)\mathbf{1}_{[0,1)}(x) + \int_0^1 u'(x-u)du,$$

where in the last equality we have used that $u'(x) = 0$, for $x < 0$. The last term may be interpreted as the convolution with the uniform distribution $\int_0^1 u'(x-u)du = \mathbb{E}(u'(x-U))$, with U being uniformly distributed in $[0,1]$.
(iii) When $\beta > 1$, by an iterative procedure and the introduction of the r.v. ν, we obtain the integral representation (24). In particular, the martingale property satisfied by the scale function $\bar{u}(m) = \mathbb{P}(\bar{U}_\infty \le m)$ implies that $\bar{u}$ is solution of the delayed equation, with $\bar{u}(0) = \mathbb{P}(\bar{U}_\infty = 0) = 1 - 1/\beta$ equal to $\mathbb{P}(S_\nu = 0)$. Thus, both distribution functions $\bar{u}(x)$ and $\mathbb{P}(S_\nu \le x)$ are solutions of the same DDE, with the same initial condition. So, they are identical, i.e. $\mathbb{P}(\bar{U}_\infty \le m) = \mathbb{P}(S_\nu \le m)$. Thanks to the first part of the theorem, the function $\tilde{u}(x) = \rho^x\bar{u}(x)$ is solution of the DDE with coefficient $\tilde{\beta}$. For the case $\beta < 1$, we can invert the role of u and $\tilde{u}$. □

We now complete the identification between the solution of the delayed equation and the scale function, defined as the function whose Laplace transform is the inverse of the Laplace exponent of the Lévy process (Bertoin [9]).

Corollary 1 (Laplace transform and scale functions)**.** *Assume* $\beta(\rho) > 1$.
Let $\psi^\rho(\alpha) = \alpha\lambda(\beta(\rho) - \beta(e^{-\alpha}))$ *be the Laplace exponent of the Lévy process* $(\mathbb{P}, U^\rho)$.
(i) *Then, the Laplace transform of* $\frac{1}{\lambda(\beta-1)}\bar{u}(x)$ *satisfies*

$$\frac{1}{\lambda(\beta-1)}\int_0^\infty e^{-\alpha x}\bar{u}(x)dx = \frac{1}{\alpha\lambda(\beta-1)}\mathbb{E}(e^{-\alpha S_\nu}) = \big(\alpha\lambda(\beta(\rho) - \beta(e^{-\alpha}))\big)^{-1}. \tag{25}$$

Then, $\frac{1}{\lambda(\beta-1)}\bar{u}(x) = 1/\psi^\rho(\alpha)$ *is the inverse of the Laplace exponent of* U^ρ, *and so by definition the scale function* $W(x,\lambda)$ *of* $(\mathbb{P}, U^\rho)$ *in the sense of Lévy processes.*
(ii) *The Laplace exponent of the Lévy process* $(\widetilde{\mathbb{P}}, U^\rho)$ *is* $\psi^\rho(\alpha - \log(\rho))$, *which is the inverse of the Laplace transform of the function* $\frac{1}{\lambda(\beta-1)}\rho^x\,\bar{u}(x)$. *So,* $\widetilde{W}(x,\rho\lambda) = \rho^x W(x,\lambda)$ *is the scale function of the Lévy process* $(\widetilde{\mathbb{P}}, U^\rho) = (\widetilde{\mathbb{P}}, \tilde{U}^{\tilde{\rho}})$.

Proof. (i) In Proposition 2, we calculated the Laplace exponent of the Lévy process U^ρ as $\psi^\rho(\alpha) = \alpha\lambda(\beta(\rho) - \beta(e^{-\alpha}))$, where λ is the intensity of the Poisson process N.

Since $\beta > 1$, by independence, the Laplace transform of S_ν is easy to calculate, and

$$\frac{1}{\alpha\lambda(\beta-1)}\mathbb{E}(e^{-\alpha S_\nu}) = \frac{1-1/\beta}{\alpha\lambda(\beta-1)}\sum_{n=0}^{\infty}(\beta(e^{-\alpha})/\beta)^n = \big(\alpha\lambda(\beta-\beta(e^{-\alpha}))\big)^{-1} = \psi^\rho(\alpha)$$

The function $W(x)$, solution of the DDE with $\beta W(0) = 1$ is the scale function of U^ρ when the intensity of N is 1, and $W(x)/\lambda$ is the scale function when the intensity is λ.
(ii) Assume $\beta > 1$, and introduce the $\widetilde{\mathbb{P}}$-Laplace exponent of the U^ρ,

$$\tilde{\psi}^\rho(\alpha) = \int_0^\infty e^{-\alpha x}\rho^x W(x)dx = \psi^\rho(\alpha - \log(\rho)), \quad \alpha > \log(\rho).$$

By similar transformations as in Subsubsection 3.1.1, $\psi^\rho(\alpha-\log(\rho))$ is the $\widetilde{\mathbb{P}}$-Laplace exponent of U^ρ, since $\rho^{U^\rho_1}$ is the density of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. $\rho^x\, W(x)$ is the scale function of $(\widetilde{\mathbb{P}}, U^\rho)$, or also the scale function $(\widetilde{\mathbb{P}}, U^{\tilde{\rho}})$. □

Other computations related to delayed differential equations associated with the performance functions are performed at the end of this section, with different boundary conditions and drifts. The solutions are only depending on the scale function and their primitives.

5.2. *Stochastic differential calculus.*

5.2.1. *Generic calculation with discontinuity.* Let us consider a generic finite variation process Z driven by a Poisson process N, and solution of the stochastic equation,

$$dZ_t = \sigma(Z_{t-})dN_t + b(Z_t)d\Lambda_t \tag{26}$$

For instance, in the family of such processes, we find the processes U $(\sigma(x) = 1, b(x) = -\beta)$ and X $(\sigma(x) = -1, b(x) = \beta)$, as well as the processes V $(\sigma(x) = \mathbf{1}_{\{x\geq 0\}}, b(x) = -\beta\mathbf{1}_{\{x>0\}}$, see (16)) and Y ($\sigma(x) = -\mathbf{1}_{\{x\geq 0\}}j(x), b(x) = \beta\mathbf{1}_{\{x\geq 0\}}$, see (17).)
Let us consider a continuous finite variation function ϕ as above. By composition with the finite variation process Z_t, the process $\phi(Z_t)$ is still of finite variation in time and

$$d\phi(Z_t) = (\phi(Z_{t-}+\sigma(Z_{t-})) - \phi(Z_{t-}))\, dN_t + \phi'(Z_t)\, b(Z_t)\, d\Lambda_t. \tag{27}$$

Moreover, when ϕ has only one discontinuity at m, the process $\phi(Z_t)$ has additional jumps due to ϕ, when $Z_t = Z_{t-} = m$ and $\delta\phi_m = \phi(m) - \phi(m^-) \neq 0$. Since the number of continuous crossings at the level m by Z is discrete, we denote this process $J^{m,Z}_t = \sum_{s\leq t}\mathbf{1}_{\{Z_s = Z_{s-} = m\}}$. So, we have to add to the previous formula the term $\delta\phi(m)\, J^{m,Z}_t$.

Theorem 6. *Let $J^{m,Z}$ be the number of continuous crossings of m, i.e. $J^{m,Z}_t = \sum_{s\leq t}\mathbf{1}_{\{Z_s=Z_{s-}=m\}}$. Recall that $M_t = N_t - \Lambda_t$ is a $\mathbb{P}_x$-local martingale.*
(i) *If ϕ is a.e. differentiable, continuous except at the level m with jump $\delta\phi(m) = \phi(m) - \phi(m^-)$, put $\Delta^\sigma\phi(x) = \phi(x+\sigma(x)) - \phi(x)$. Then, for any function α (positive or bounded), the differential formula becomes,*

$$\begin{aligned} d\phi(Z_t) + \delta\phi(m)dJ^{m,Z}_t + \alpha(Z_{s-})dN_s &= \Delta^\sigma\phi(Z_{t-})dN_t + \phi'(Z_t)\, b(Z_t)\, d\Lambda_t, \\ &= -\phi'(Z_{t-})b(Z_{t-})dM_t + (\Delta^\sigma\phi + \phi'\, b + \alpha)(Z_{t-})dN_t. \end{aligned} \tag{28}$$

(ii) *The left-hand side is a local martingale if and only if ϕ is solution of the delayed equation*

$$\Delta^\sigma\phi(z) + \phi'(z)\, b(z) + \alpha(z) = 0, \quad a.e. \tag{29}$$

In particular, since the process $\bar{u}(m-U^{\rho}_t)$ is a martingale in the domain $(0,m)$, the last condition implies after the change of variable $x = m - z$, that for $x > 0$

$$\bar{u}'(x)\beta = \bar{u}(x) - \bar{u}(x-1) \quad a.e..$$

The applications to the performance functions are of two types. On one hand, they allow us to extend the martingale property of the processes H^m and G^m defined in (18) at any time. On the other hand, they allow us to precise the Neumann delayed equation satisfied by the performance functions. As before, any results can be given on the same probability space. We frequently use the notation $N^{m,Z}_t = \int_0^t \mathbf{1}_{[0,m]}(Z_{s-})dN_s$ and the definition of the performance functions given in Section 4.1.3.

Proposition 7. (i) *Let us consider the process V solution of the differential equation* (16)*.*

a) *The process $H^m_t = h_m(V_t) - h_m(V_0) + N^m_t - h_m(m-)J^{m,d}_t$ is the $\mathbb{P}_x$- local martingale*

$$dh_m(V_t) + dN^m_t - h_m(m-)dJ^{m,d}_t = dH^m_t = \beta h'_m(V_{t-})dM_t. \tag{30}$$

b) *So, h_m is the Neumann solution with jump at m, of the delayed equation*

$$\beta h'_m(x) = h_m(x+1) - h_m(x) + 1, \quad x \in (0,m), \quad h'(0) = 0, \quad h(x) = 0,\ x \geq m. \tag{31}$$

(ii) *Let us now consider the process Y_t solution of the differential equation* (17)*. The function g_m is continuous at m.*

a) *The process $G^m_t = g_m(Y_t) - g_m(Y_0) + N^{m,Y}_t$ is the following $\mathbb{P}_x$- local martingale ,*

$$dg_m(Y_t) + dN^{m,Y}_t = dG^m_t = \beta g'_m(Y_{t-})dM_t. \tag{32}$$

b) *So, g_m is the continuous solution of the Cauchy delayed equation,*

$$\beta g'_m(x) = g_m(x) - g_m((x-1)^+) - 1, \quad x \in (0,m), \quad g_m(x) = 0,\ x \geq m. \tag{33}$$

Proof. The proof is immediate from Theorem 6, and the differential equations of V and Y. The same delayed equations (31) and (33) hold true for the tilded-functions using the parameter $\tilde{\beta} = \beta(1/\rho) = \beta(\rho)/\rho$ instead of β. The resolution of the delayed equations is postponed to the next subsection. $\square$

5.2.2. *Computation of the performance functions with the help of scale functions.* The performance functions $h_m(x) = \mathbb{E}_x(N_{\tau^V_m})$ and $g_m(x) = \mathbb{E}_x(N_{\tau^Y_m})$ are solutions of the DDEs (31) and (33) respectively, with different boundary conditions. Thanks to Theorem 5, the functions $h_m(m-z) = \mathsf{k}_m(z)$ and $g_m(y)$ are linear combinations of the scale functions W and their primitive $\widehat{W}(z) = \int_0^z W(y)dy$ (without restriction on the value of the parameter $\beta(\rho)$). The functions W and $\widetilde{W}$ are defined by 0 for $x < 0$ and

$$\begin{cases} \text{for } \rho > 1, \quad W(x) = \dfrac{1}{(\beta-1)}\mathbb{P}(\bar{U}^{\rho}_{\infty} \leq x), \quad \widetilde{W}(x) = \rho^x\, W(x), \\ \text{for } \ \rho < 1, \quad W(x) = \rho^{-x}\widetilde{W}(x), \qquad\qquad \widetilde{W}(x) = \dfrac{1}{\rho(\tilde{\beta}-1)}\,\widetilde{\mathbb{P}}(\bar{U}^{\rho}_{\infty} \leq x). \end{cases} \tag{34}$$

The following results are not new, and similar closed formulae may be derived directly from [28]. Nevertheless, in the Poisson case, our proofs are very elementary and do not appeal to the excursion theory. Given our simplified framework, we only need elementary differential calculus.

Theorem 8 (Closed formulas)**.** (i) *The performance functions $g_m(y)$ and $\tilde{g}_m(y)$, extended by* 0 *for $y < 0$, are continuous solutions on $(0, m)$, null for $y > m$ of the delayed equations,*

$$\beta\, g'_m(y) \;=\; g_m(y) - g_m(y-1) - 1, \quad \tilde{\beta}\tilde{g}'_m(y) = \tilde{g}_m(y) - \tilde{g}_m(y-1) - 1, \tag{35}$$

$$\text{Then,} \quad g_m(y) \;=\; \int_y^m W(z)dz, \quad \tilde{g}_m(y) = \int_y^m \rho\widetilde{W}(z)dz, \quad y \in [0, m]. \tag{36}$$

(ii) *The modified performance functions $\mathsf{k}_m(z) = h_m(m-z)$, and $\tilde{\mathsf{k}}_m(z) = \tilde{h}_m(m-z)$, extended by* 0 *for $z < 0$, are solutions on $(0, m)$ of the delayed equations* (35),

$$\beta\, \mathsf{k}'_m(y) = \mathsf{k}_m(y) - \mathsf{k}_m(y-1) - 1, \quad \tilde{\beta}\tilde{\mathsf{k}}'_m(y) = \tilde{\mathsf{k}}_m(y) - \tilde{\mathsf{k}}_m(y-1) - 1. \tag{37}$$

satisfying the Neumann condition $\mathsf{k}'_m(m-) = 0$, $\tilde{\mathsf{k}}'_m(m-) = 0$.
Then, k_m and $\tilde{\mathsf{k}}_m$ are linear combinations of $(W, \widehat{W})$, or $(\widetilde{W}, \widehat{\widetilde{W}})$, and

$$h_m(x) \;=\; W(m-x)\frac{W(m)}{W'(m)} - \int_0^{m-x} W(y)dy, \quad h_m(m-) = W(0)\frac{W(m)}{W'(m)}, \tag{38}$$

$$\tilde{h}_m(x) \;=\; \rho\Big(\widetilde{W}(m-x)\frac{\widetilde{W}(m)}{\widetilde{W}'(m)} - \int_0^{m-x} \widetilde{W}(y)dy\Big), \quad \tilde{h}_m(m-) = \rho W(0)\frac{\widetilde{W}(m)}{\widetilde{W}'(m)}. \tag{39}$$

Proof. (i) Recall that the primitive of any solution u of the delayed equation is solution of the delayed equation (35) with drift $\beta u(0)$. When the drift is one, the reference solution is the scale function W since $W(0) = 1/\beta$. When the drift is -1, the reference solution is $-W(y)$. Associated with the boundary condition $g_m(m) = 0$, we obtain the relation (36). Similarly, the same relation holds true for the tilded equation, with the drift condition $\tilde{\beta}\tilde{u}(0) = -1$. So, the reference solution is $-\rho\widetilde{W}(y)$ whose value at 0 is $-\rho W(0) = -1/\tilde{\beta}$.
(ii) The function $\mathsf{k}_m(z)$ is solution of the linear delayed equation with drift -1, and Neumann condition $\mathsf{k}'_m(0) = 0$. So, we are looking for solutions which are linear combinations of W and $\widehat{W}$. Since the constant is equal to -1, the coefficient of $\widehat{W}$ is given by -1. However, we need an additional term proportional to $W(z)$ to satisfy the Neumann condition $\mathsf{k}'_m(m-) = 0$. Therefore, we are looking for a function $\mathsf{k}_m(z) = \alpha_m W(z) - \widehat{W}(z)$, with left-derivative at m equal to 0, i.e. $\alpha_m W'(m-) - W(m) = 0$. The formula (38) gives an explicit form at the relation $h_m(x) = \mathsf{k}_m(z) = \alpha_m W(m-x) - \widehat{W}(m-x)$.

(iii) For the tilded equation, just as for $\tilde{g}_m$, the coefficient of $\widehat{\widetilde{W}}$ is $-\rho$. We are also looking for a coefficient $\tilde{\alpha}_m$ such that $\tilde{\alpha}_m\widetilde{W}'(m-) - \rho\widetilde{W}(m) = 0$. So, the closed formula (39) is based on this observation.
(iii) Assume now a λ-intensity and a scale function $W(x, \lambda) = (1/\lambda)W(x)$. So the solution of the delayed equation when the scale function is $W(x, \lambda)$ is $\lambda W(x, \lambda) = W(x)$, and yield to the invariance by scaling of the performance function $g_m(x) = \mathbb{E}_x(N_{\tau_m^Y})$. The same argument holds true for the other performance functions. $\square$

6. Optimality of the cusum stopping rule.

Since the cusum process depends on the value of ρ, we reintroduce the distinction between $\rho > 1$ associated with the process V and the performance functions h_m and $\tilde{h}_m$ and $\rho < 1$ associated

with the process Y and the performance functions g_m and $\tilde{g}_m$. The aim is to prove the optimality of the stopping time τ_m^V (resp. τ_m^Y) in the family of finite stopping times T with finite cusum performance, i.e. $C(T) = \sup_{\theta\in[0,\infty]} \operatorname{ess\,sup}_\omega \ \widetilde{\mathbb{E}}\,((N_T - N_\theta)^+|\mathcal{F}_\theta)$, and satisfying the false alarm constraint: $\mathbb{E}[N_T] \geq \mathbb{E}[N_{\tau_m^V}] = \mathbb{E}[N^m_{\tau_m}] = h_m(0)$, (resp. $\mathbb{E}[N_T] \geq \mathbb{E}[N^m_{\tau_m^Y}] = g_m(0)$).
The first step is to provide an equivalent criterion which allows us to work only with the probability measure $\mathbb{P}$.

6.1. *Cusum criterion and its modification.* The value of the cusum criterion associated with a cusum stopping rule τ_m^V or τ_m^Y is easy to obtain from the martingale properties established in Proposition 7 since the functions $\tilde{h}_m$ and $\tilde{g}_m$ are decreasing. But showing its optimality is more complex. In order to take the false alarm constraint into account we work under the same probability measure $\mathbb{P}$. Therefore, in the same vein as Shiryaev [34] and Moustakides [24], we introduce a modification of the Lorden criterion, providing a convenient lower bound for the conditional worst case performance, still reached by the same cusum stopping times. A useful tool is an integration by parts formula under the probability $\widetilde{\mathbb{P}}_x$.

Proposition 9. *Let us consider the process $\Gamma_t^T := \widetilde{\mathbb{E}}_x(\int_t^T dN_s \big| \mathcal{F}_t)$. Let $(\bar{Z}_t)$ be a càdlàg monotonic adapted process whose jumps occur only at the jump epochs of N as $\bar{X}_t$ or $-\bar{U}_t$. Then, we have*

$$\widetilde{\mathbb{E}}_x\Big[\int_t^T \Gamma_\alpha^T d\rho^{\bar{Z}_\alpha}\big|\mathcal{F}_t\Big] = \widetilde{\mathbb{E}}_x\Big[\int_t^T (\rho^{\bar{Z}_{s-}} - \rho^{\bar{Z}_t}) dN_s\big|\mathcal{F}_t\Big]. \tag{40}$$

Several important consequences under the probability measures $\mathbb{P}_x$ can be made explicit.

(i) *When $(\bar{Z} = \bar{X}^{\mathrm{ad}}, T = \tau_m^V := \tau_m)$, or $(\bar{Z} = -\bar{U}^{\mathrm{ad}}, T = \tau_m^Y := \tau_m)$, we have*

$$\rho^x(\tilde{h}_m(x) - \tilde{h}_m(0)) = \rho\,\mathbb{E}_x\Big(\int_0^{\tau_m} \rho^{V_{s-}} dN_s\Big) - \tilde{h}_m(0)\,\mathbb{E}_x(\rho^{V_{\tau_m}}), \tag{41}$$

$$\rho^{-y}(\tilde{g}_m(y) - \tilde{g}_m(0)) = \rho\,\mathbb{E}_y\Big(\int_0^{\tau_m} \rho^{-Y_{s-}} dN_s\Big) - \tilde{g}_m(0)\,\mathbb{E}_y(\rho^{-Y_{\tau_m}}). \tag{42}$$

a) *So, $H_t^\rho = \int_0^t \rho^{V_s} dN_s^{m,V} - \rho^m \tilde{h}_m(m-) J_t^{m,V} + \rho^{V_t}(\tilde{h}_m(Y_t) - \tilde{h}_m(0))$ is a $\mathbb{P}_x$-martingale*

b) *and $G_t^{\rho,m} = \rho \int_0^t \rho^{-Y_{s-}} dN_s^{m,Y} + \rho^{-Y_t}(\tilde{g}_m(Y_t) - \tilde{g}_m(0))$ is a $\mathbb{P}_x$-martingale.*

(ii) *When Γ_t^T is bounded by $C(T)$, the following lower bounds hold:*

$$\text{for } \rho > 1, \qquad \rho\,\mathbb{E}\Big[\int_t^T \rho^{V_{s-}} dN_s\big|\mathcal{F}_t\Big] \leq C(T)\,\widetilde{\mathbb{E}}\big(\rho^{V_T}\big|\mathcal{F}_t\big), \tag{43}$$

$$\text{for } \rho < 1, \qquad \rho\,\mathbb{E}\Big[\int_t^T \rho^{-Y_{s-}} dN_s\big|\mathcal{F}_t\Big] \leq C(T)\,\mathbb{E}\big(\rho^{-Y_T}\big|\mathcal{F}_t\big). \tag{44}$$

(iii) *$\tilde{h}_m(0)$ and $\tilde{g}_m(0)$ are respectively the cusum bounds of the stopping times τ_m^V ($\rho > 1$) and τ_m^Y ($\rho < 1$).*

Proof. (i) We start with calculation under the probability $\widetilde{\mathbb{P}}_x$ and take the primitive of Γ_t^T with respect to the increasing process $\rho^{\bar{Z}}$. To prove the equalities, it is equivalent to introduce any stopping times $S \leq T$, and work with the expectations in place of conditional expectations. Using integration by parts formula, we show that (40) is equivalent to

$$\begin{aligned}\widetilde{\mathbb{E}}\Big[\int_S^T \Gamma_\alpha^T d\rho^{\bar{Z}_\alpha}\Big] &= \widetilde{\mathbb{E}}\Big[\int_S^T \Big(\int_{(\alpha,T]} dN_u\Big) d\rho^{\bar{Z}_\alpha}\Big], \\ &= \widetilde{\mathbb{E}}\Big[\int_S^T dN_u \Big(\int_{(S,u)} d\rho^{\bar{Z}_\alpha}\Big)\Big] = \widetilde{\mathbb{E}}\Big[\int_S^T dN_u (\rho^{\bar{Z}_{u-}} - \rho^{\bar{Z}_S})\Big].\end{aligned}$$

(ii) The first application is the comparison of the performance functions.
a) Starting with $T = \tau^V_m =: \tau_m$ and $\bar Z = \bar X^{\rm ad}$, the process Γ^T_t is a function of V_t, $\Gamma^T_t = \tilde h_m(V_t)$. Since $\rho^{\bar X^{\rm ad}}_t$ only increases when $V = 0$, the equality (40) at $t = 0$ for $x \le m$ becomes

$$\tilde h_m(0)\, \widetilde{\mathbb E}_x(\rho^{\bar X^{\rm ad}_{\tau_m}} - 1) = \widetilde{\mathbb E}_x(\int_0^{\tau_m} (\rho^{\bar X^{\rm ad}_{s-}} - 1) dN_s) = \widetilde{\mathbb E}_x(\int_0^{\tau_m} \rho^{\bar X^{\rm ad}_{s-}} dN_s) - \tilde h_m(x).$$

So, $\widetilde{\mathbb E}_x(\int_0^{\tau_m} \rho^{\bar X^{\rm ad}_s} dN_s) - \tilde h_m(0)\widetilde{\mathbb E}_x(\rho^{\bar X^{\rm ad}_{\tau_m}}) = \tilde h_m(x) - \tilde h_m(0)$.
b) This form is well-adapted to pass from $\widetilde{\mathbb P}_x$ to $\mathbb P_x$ since $\widetilde{\mathbb P}_x/\mathbb P_x = \rho^{U_{\tau_m} - U_0}$. Using that $\rho^{U_t - U_0}$ is a density martingale, together with the relation $U_t(U_0) + \bar X^{\rm ad}_t(U_0) = V_t(U_0)$, we can replace $\bar X^{\rm ad}$ by $\bar V^{\rm ad} - U_0$ in the previous equality to obtain the result under $\mathbb P_x$; in other words,

$$\rho^{-x}\big(\mathbb E_x(\int_0^{\tau_m} \rho^{V_s} dN_s) - \tilde h_m(0)\mathbb E_x(\rho^{V_{\tau_m}})\big) = \tilde h_m(x) - \tilde h_m(0).$$

(ii) a) The same argument can be used when starting with $T = \tau^Y_m =: \tau_m$ and $\bar Z = -\bar U^{\rm ad}$, since $\Gamma^T_t = \tilde g_m(Y_t)$. We have to pay attention to the discontinuities of $\bar U^{\rm ad}$, but not in the left-hand side of the equality since the support of $\bar U^{\rm ad}$ is the set $\{Y_t = 0\}$. Then, we still have

$$\widetilde{\mathbb E}_y(\int_0^{\tau_m} \rho^{-\bar U^{\rm ad}_{s-}} dN_s) - \tilde g_m(0)\widetilde{\mathbb E}_y(\rho^{-\bar U^{\rm ad}_{\tau_m}}) = \tilde g_m(y) - \tilde g_m(0).$$

We also have that $\rho^{U_t - U_0}$ is a $\mathbb P_y$-martingale, and $U_t(U_0) - \bar U^{\rm ad}_t(U_0) = U_0 - Y_t(U_0)$. So by identification, we obtain that $\widetilde{\mathbb E}_y(\rho^{-\bar U^{\rm ad}_{\tau_m}}) = \mathbb E_y(\rho^{U_{\tau_m} - \bar U^{\rm ad}_{\tau_m}}) = \rho^y\, \mathbb E_y(\rho^{-Y_{\tau_m}})$.
b) The other term involves the left limit of $\bar U^{\rm ad}$, so we use the identity $U_t - \bar U^{\rm ad}_{t-}(y) = y - Y_{t-}(y) + (U_t - U_{t-})$. Since $U_t - U_{t-} = 1$ at any jump dates of N, we have

$$\widetilde{\mathbb E}_y(\int_t^{\tau_m} \rho^{-\bar U^{\rm ad}_{s-}} dN_s) = \mathbb E_y(\int_t^{\tau_m} \rho^{U_s - \bar U^{\rm ad}_{s-}} dN_s) = \rho^y\, \mathbb E_y(\int_t^{\tau_m} \rho\, \rho^{-Y_{s-}} dN_s).$$

So as before,

$$\rho^y\Big(\mathbb E_y(\int_0^{\tau_m} \rho\, \rho^{-Y_{s-}} dN_s) - \tilde g_m(0)\, \mathbb E_y(\rho^{-Y_{\tau_m}})\Big) = \tilde g_m(x) - \tilde g_m(0).$$

This formula differs from the one involving the process V by the fact that dN_t a.e. $V_t - V_{t-} = 1$ but $-(Y_t - Y_{t-}) \neq 1$.
(iii) a) The last application is relative to the cusum bounds. If Γ^T_t is bounded by $C(T)$, and $\rho^{\bar Z_t}$ is non-decreasing, the left-hand side of (40) is dominated by $C(T)\widetilde{\mathbb E}\big[\rho^{\bar Z_T} - \rho^{\bar Z_t}\big|\mathcal F_t\big]$.
From the right-hand side, we deduce that $\widetilde{\mathbb E}\big[\int_t^T (\rho^{\bar Z_{s-}} - \rho^{\bar Z_t}) dN_s\big|\mathcal F_t\big] \le C(T)\widetilde{\mathbb E}\big[\rho^{\bar Z_T} - \rho^{\bar Z_t}\big|\mathcal F_t\big]$.
Since $\rho^{\bar Z_t}$ is $\mathcal F_t$-measurable, and $\rho^{\bar Z_t}\Gamma^T_t \le C(T)\rho^{\bar Z_t}$, we also have the simplified relation $\widetilde{\mathbb E}\big[\int_t^T \rho^{\bar Z_{s-}} dN_s\big|\mathcal F_t\big] \le C(T)\widetilde{\mathbb E}\big(\rho^{\bar Z_T}\big|\mathcal F_t\big)$.
b) The monotony property is verified for instance if $(\rho > 1, \bar X)$ or $(\rho < 1, -\bar U)$. Assume $\rho > 1$ and $\bar Z = \bar X$. As previously, we can transform this inequality as a $\mathbb P$-inequality, since the conditional density of $\widetilde{\mathbb P}$ with respect to $\mathbb P$ is ρ^{U_T}/ρ^{U_t}, so that as in the case of functions, $\mathbb E\big[\int_t^T \rho^{V_s} dN_s\big|\mathcal F_t\big] \le C(T)\mathbb E\big(\rho^{V_T}\big|\mathcal F_t\big)$.
Assume $\rho < 1$ and $\bar Z = -\bar U$. Using the same argument as in the case of functions, we obtain $\mathbb E\big[\int_t^T \rho^{(1 - Y_{s-})} dN_s\big|\mathcal F_t\big] \le C(T)\mathbb E\big(\rho^{-Y_T}\big|\mathcal F_t\big)$.
c) When $T = \tau_m$, at time 0, we have $\mathbb E(\int_0^{\tau_m} \rho^{V_s} dN_s) = \tilde h_m(0)\mathbb E(\rho^{Y_{\tau_m}})$, and
$\mathbb E(\int_0^{\tau_m} \rho^{1 - Y_{s-}} dN_s) = \tilde g_m(0)\mathbb E(\rho^{-V_{\tau_m}})$. It thus follows that $\tilde h_m(0)$ and $\tilde g_m(0)$ are respectively the cusum bounds of the stopping times τ^V_m $(\rho > 1)$ and τ^Y_m $(\rho < 1)$. □

6.2. *Optimality*

(i) False alarm constraint. The optimality of the cusum stopping time τ_m is to be restricted to the class of stopping times with finite cusum performance and satisfying the false alarm constraint, $\mathbb{E}(N_{\tau_m}) \leq \mathbb{E}(N_T) < +\infty$. So, it is sufficient to prove that

$$\mathbb{E}\big(\int_0^T \rho^{V_s} dN_s\big) \geq \tilde{h}_m(0)\,\mathbb{E}\big(\rho^{V_T}\big),\ \text{given}\ \mathbb{E}(N_T) = \mathbb{E}(N_{\tau_m}) = h_m(0). \tag{45}$$

Proposition 7 provides useful tools to reformulate these inequalities in a more tractable form, by introducing the V-processes $N_t^{m,V} = \int_0^t \mathbf{1}_{[0,m)}(V_{s-})dN_s$ and $\int_0^t \rho^{V_s} dN_s^m$, or the Y-processes $N_t^{m,Y} = \int_0^t \mathbf{1}_{[0,m)}(Y_{s-})dN_s$ and $\int_0^t \rho^{1-Y_{s-}} dN_s^m$. We first start with the case $\rho < 1$, being the simplest from the optimality point of view.

Theorem 10 (Optimality result for a decrease in the intensity)**.** *Assume $\rho < 1$, and consider cusum process Y. Let T be a stopping time with finite cusum performance, and false alarm constraint $\mathbb{E}(N_{\tau_m^Y}) = \mathbb{E}(N_T) < +\infty$.*

(i) *The function $\psi(y) = \rho^{-(m-y)} g_m(y) - \tilde{g}_m(y)/\rho$, defined on $[0, m]$ is positive.*

(ii) *$\tilde{g}_m(0)$ is a lower bound for the criterion*

$$\rho\mathbb{E}\big(\int_0^T \rho^{-Y_{s-}} dN_s\big)/\mathbb{E}\big(\rho^{-Y_T}\big) \geq \tilde{g}_m(0),\ \textit{given}\ \mathbb{E}(N_T) = \mathbb{E}(N_{\tau_m^Y}) = g_m(0). \tag{46}$$

(iii) *τ_m^Y is an optimal detection rule for the problem (4) under the false alarm constraint $\mathbb{E}(N_T) = \mathbb{E}(N_{\tau_m^Y}) = g_m(0)$.*

Proof. Assume $\rho < 1$. The cusum process is the process Y, with cusum stopping time τ_m^Y (τ_m in short). Recall that $N_t^{m,Y} = \int_0^t \mathbf{1}_{[0,m)}(Y_{s-})dN_s$.

(i) Indeed, by Proposition 9, and from the continuity of the function g_m, for any stopping time S such that $\mathbb{E}(N_S^{m,Y}) < \infty$, we have $g_m(0) = \mathbb{E}(N_S^{m,Y} + g_m(Y_S))$. The false alarm constraint becomes:

$$\mathbb{E}(\textstyle\int_0^T \mathbf{1}_{[m,\infty)}(Y_{s-})dN_s) = \mathbb{E}(g_m(Y_T)).$$

By Proposition 9, we know that $\int_0^t \rho^{1-Y_{s-}} dN_s + \rho^{-Y_t}(\tilde{g}_m(Y_t) - \tilde{g}_m(0))$ is a martingale, null at time 0, so that

$$\mathbb{E}\big(\int_0^T \rho^{1-Y_{s-}} dN_s^m - \tilde{g}_m(0)\rho^{-Y_T} + \rho^{-Y_T}\tilde{g}_m(Y_T)\big) = 0.$$

Therefore, we have to show that, for $\rho < 1$,

$$\mathbb{E}\big(\int_0^T \rho^{1-Y_{s-}} \mathbf{1}_{[m,\infty)}(Y_{s-})dN_s - \rho^{-Y_T}\tilde{g}_m(Y_T)\big) \geq 0,$$

if $\mathbb{E}(\int_0^T \mathbf{1}_{(m,\infty)}(Y_{s-})dN_s - g_m(Y_T)) = 0$. The idea is to control the term $\rho^{1-Y_{s-}}\mathbf{1}_{[m,\infty)}(Y_{s-})$ by $\rho^{1-m}\mathbf{1}_{[m,\infty)}(Y_{s-})$ since the difference is still nonnegative ($\rho < 1$). So, the inequality will be proved if we show that $\rho^{1-m} g_m(Y_T) - \rho^{-Y_T}\tilde{g}_m(Y_T)$ is nonnegative in expectation.

(ii) a) Let us study the function $\psi(y) = \rho^{-(m-y)} g_m(y) - \tilde{g}_m(y)/\rho$, equal to 0 when $y \geq m$, using the description of the functions g_m and $\tilde{g}_m$ given in Theorem 8 in terms of the scale functions. Recall that $g_m(y) = \int_y^m W(z)dz$ and $\tilde{g}_m(y) = \int_y^m \rho\widetilde{W}(z)dz$.

b) The derivative $\psi'(y)$ is negative since $A = \log(\rho)\rho^{-(m-y)} g_m(y)$ is negative and

$$\begin{aligned}\psi'(y) - A &= -\rho^{-(m-y)}W(y) + \rho\,\widetilde{W}(y)/\rho = \rho^y W(y) - \rho^{-(m-y)}W(y),\\ &= \rho^y W(y)(1 - \rho^{-m}) \leq 0, \quad (\rho < 1).\end{aligned}$$

So $\psi'(y)$ is negative and $\psi(y)$ is non-negative. Thus, the lower bound is verified.
(iii) We have seen in Proposition 9 that the lower bound is an equality for τ_m^Y. Thus, the optimality is proved. $\square$

The case $\rho > 1$ is more delicate, since the performance functions are discontinuous at the level m, with jumps $-h_m(m-) = -W(0)\frac{W(m)}{W'(m)}$ and $-\tilde{h}_m(m-) = -\rho W(0)\frac{\widetilde{W}(m)}{\widetilde{W}'(m)}$, where the last equalities are given in Theorem 8, together with the following formulae:

$$\begin{aligned} h_m(z) &= W(m-z)\frac{h_m(m-)}{W(0)} - \int_0^{m-z} W(y)dy, \\ \tilde{h}_m(z) &= \rho^{m-z}W(m-x)\frac{\tilde{h}_m(m-)}{W(0)} - \int_0^{m-z} \rho^y W(y)dy. \end{aligned}$$

Theorem 11 (Optimality result for an increase in the intensity)**.** *Assume $\rho > 1$, and consider the cusum process V. Let T be a stopping time with finite cusum performance, and false alarm constraint $\mathbb{E}(N_{\tau_m^V}) = \mathbb{E}(N_T) < +\infty$.*
(i) *The function $\phi_m(z)$ is continuous and non-negative on $[0, m)$, where*

$$\phi_m(m-z) = \frac{\tilde{h}_m(m-)}{h_m(m-)}\,\rho^{m-z}\,h_m(z) - \tilde{h}_m(z) = \rho\int_0^{m-z}\widetilde{W}(y)dy - \frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^{m-z}\int_0^{m-z} W(y)dy. \quad (47)$$

(ii) *$\tilde{h}_m(0)$ is a lower bound for the criterion*

$$\mathbb{E}\big(\int_0^T \rho^{V_s}dN_s\big)/\mathbb{E}\big(\rho^{V_T}\big) \geq \tilde{h}_m(0),\ \textit{given}\ \mathbb{E}(N_T) = \mathbb{E}(N_{\tau_m^V}) = h_m(0). \quad (48)$$

(iii) *τ_m^V is an optimal detection time for the false alarm constraint $\mathbb{E}(N_T) = \mathbb{E}(N_{\tau_m^V}) = h_m(0)$.*

Proof. Recall that $N_t^{m,V} = \int_0^t \mathbf{1}_{[0,m)}(V_{s-})dN_s$ and $J_{\cdot}^{d,m}$ is the number of continuous down crossings of m by V.
(i) By Proposition 7, since the decreasing function h_m has a jump of $-h_m(m-)$ at m, for any stopping time T, we have: $h_m(0) = \mathbb{E}(N_T^m + h_m(V_T) - h_m(m-)\,J_T^{d,m}$.
Similarly, the false alarm constraint becomes:

$$\mathbb{E}(\textstyle\int_0^T \mathbf{1}_{[m,\infty)}(V_{s-})dN_s + h_m(m-)\,J_T^{d,m}) = \mathbb{E}(h_m(V_T)).$$

From Proposition 9, and the martingale property, we know that

$$\mathbb{E}\big(\textstyle\int_0^T \rho^{V_s}dN_s^{m,V} - \tilde{h}_m(0)\rho^{V_T}\big) = \mathbb{E}\big(\tilde{h}_m(m-)\rho^m J_T^{d,m} - \rho^{V_T}\tilde{h}_m(V_T)\big).$$

So, we have to show that for $\rho > 1$,

$$\mathbb{E}\big(\textstyle\int_0^T \rho^{V_s}\mathbf{1}_{[m,\infty)}(V_{s-})dN_s + \tilde{h}_m(m-)\rho^m J_T^{d,m} - \rho^{V_T}\tilde{h}_m(V_T)\big) \geq 0,$$

given that $\mathbb{E}(\int_0^T \mathbf{1}_{(m,\infty)}(V_{s-})dN_s + h_m(m-)\,J_T^{d,m}) = \mathbb{E}(h_m(V_T))$.

(ii) Since, only a limited information is accessible about the variable $J_T^{d,m}$, we modify the criterion by multiplying the constraint by $\frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^m$, and by making the difference. The inequality becomes

$$\mathbb{E}\big(\int_0^T (\rho^{V_s} - \frac{\tilde{h}_m(m-)}{h_m(m-)})\rho^m\mathbf{1}_{[m,\infty)}(V_{s-})dN_s + \frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^m\,h_m(V_T) - \rho^{V_T}\tilde{h}_m(V_T)\big) \geq 0.$$

In fact, this inequality is true pathwise, since $\tilde{h}_m(m-) \leq h_m(m-)$,

a) Since $\frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^m \leq \rho^m$, the first term is obviously non-negative since $V_s \geq m+1$.

b) The second term is explained by the properties of the function $\phi_m(z)$, where $\phi_m(m-z) = \frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^{m-z}h_m(z) - \tilde{h}_m(z)$, whose left limit at m is equal to 0 and so the function ϕ_m is continuous. By the definition of the functions h_m and $\tilde{h}_m$, it is easy to verify that the contribution of the functions $W(m-x)$ and $\widetilde{W}(m-x)$ is canceled, and

$$\phi_m(x) = \rho \int_0^x \widetilde{W}(ydy) - \frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^x \int_0^x W(y)dy, \phi(0) = 0.$$

But this representation is not well-adapted to study the sign of its derivative. Therefore, using the previous notation $\mathsf{k}_m(z) = h_m(m-z)$ and $\tilde{\mathsf{k}}_m(z) = \tilde{h}_m(m-z)$, we come back to the initial definition $\phi_m(x) = \frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^x \mathsf{k}_m(z) - \tilde{\mathsf{k}}_m(z)$ and to the following form of the derivative:

$$\phi_m'(x) = \frac{\tilde{h}_m(m-)}{h_m(m-)} \log\rho\, \rho^x \mathsf{k}_m(x) + (\frac{\tilde{h}_m(m-)}{h_m(m-)}\rho^x \mathsf{k}_m'(z) - \tilde{\mathsf{k}}_m'(x)).$$

Since $\frac{\tilde{h}_m(m-)}{h_m(m-)} \log\rho\, \rho^{m-z}\mathsf{k}_m(z)$ is non-negative, we essentially have to study the relation between the derivatives, $B(x) = \frac{\tilde{h}_m(m-)}{\mathsf{k}_m(m-)}\, \rho^x \mathsf{k}_m'(x) - \tilde{\mathsf{k}}_m'(x)$.

c) Recall that

$$\frac{\tilde{h}_m(m-)}{W(0)} = \rho \frac{\widetilde{W}(m)}{\widetilde{W}'(m)} \qquad \text{and} \qquad \frac{\widetilde{W}'(m)}{\widetilde{W}(m)} = \log\rho + \frac{W'(m)}{W(m)} = \log\rho + \frac{W(0)}{h_m(m-)}.$$

Given that $\tilde{\mathsf{k}}_m(x) = \frac{\rho W(0)}{\tilde{h}_m(m-)}\widetilde{W}(x) - \rho\int_0^x \rho^y W(y)dy$ and $\widetilde{W}'(z) = \rho^z[\log\rho\, W(z) + W'(z)]$, it follows

$$\tilde{\mathsf{k}}_m'(x) = \rho^x W(x)\big[\frac{\tilde{h}_m(m-)}{W(0)}\log(\rho) - \rho\big] + \rho^x \frac{\tilde{h}_m(m-)}{W(0)} W'(x).$$

Between brackets, the coefficient of $-\frac{\rho W(0)}{\tilde{h}_m(m-)}\rho^x W(x)$ is

$$\frac{\rho W(0)}{\tilde{h}_m(m-)} - \log\rho = \frac{\widetilde{W}'(m)}{\widetilde{W}(m)} - \log\rho = \frac{W'(m)}{W(m)} = \frac{W(0)}{h_m(m-)}.$$

Then, after some algebra we obtain a remarkable identity on the derivatives:

$$\tilde{\mathsf{k}}_m'(x) = \rho^x \frac{\tilde{h}_m(m-)}{h_m(m-)}[-W(x) + \frac{h_m(m-)}{W(0)}W'(x)) = \rho^x \frac{\tilde{h}_m(m-)}{h_m(m-)}\, \mathsf{k}_m'(x). \tag{49}$$

In other words, the function $B(x) = \frac{\tilde{h}_m(m-)}{h_m(m-)}\, \rho^x \mathsf{k}_m'(x) - \tilde{\mathsf{k}}_m'(x)$ is the null function, and $\phi_m'(x) = \frac{\tilde{h}_m(m-)}{h_m(m-)} \log\rho\, \rho^z \mathsf{k}_m(z)$ is positive, and increasing on $(0,m)$. So, the function $\phi_m(x)$ is convex, increasing, on $(0,m)$.

(iii) Consequently, the function $\phi_m(m-z)$ is positive, decreasing on $[0,m]$ and still convex. The lower bound is established. We have seen in Proposition 9 that the lower bound is an equality for τ_m^V. The optimality is proved. □

Remark 1. These theorems give the infimum of the worst mean number of jumps until detection, i.e. $\sup_{\theta\in[0,\infty]} \operatorname*{ess\,sup}_{\omega}\ \widetilde{\mathbb{E}}\big[(N_T - N_\theta)^+ \big| \mathcal{F}_\theta\big]$ for a class of stopping times with preassigned rate of false alarm, $\mathbb{E}(N_T) \geq \pi$. The bounds depend on whether an increase or a decrease in intensity is investigated. As observed by Basseville and Nikiforov [6], this result is important not only for the cusum algorithm, but in general, since in some sense, *these lower bounds play the same role in the change detection theory as the Cramer-Rao lower bound in estimation theory.*

7. Numerical illustrations.

In this section we provide an illustration of the performance of the cusum procedure. Recall that the latter depends only on the performance functions $h_m, \tilde{h}_m, g_m$ and $\tilde{g}_m$, given in Theorem 8 in terms of the scale functions, and their primitive. Therefore, we give a closed formula for the scale function and its primitive.

7.1. *Closed formulas for the scale function.* For $\beta > 1$, recall from Equation (34) that the scale function is given as $W(x) = \mathbb{P}(S_\nu \leq x)/(\beta - 1)$.
(i) The r.v. ν has a geometric distribution with parameter $1/\beta$, i.e. $\mathbb{P}(\nu = n) = (1 - (1/\beta))\beta^{-n}$, and $S_n = \sum_{i=1}^n U_i$ the sum of n v.a U_i being i.i.d. and uniformly distributed on $(0, 1)$. The density p_u^{*n} of S_n is known as the Irwin-Hall density [16] equal to 0 when $x \geq j$ and to

$$p_u^{*n}(x) = \frac{1}{(n-1)!} \sum_{k=0}^{\lfloor x \rfloor} (-1)^k \binom{n}{k} (x-k)^{n-1}, \qquad \text{when } 0 < x < j. \tag{50}$$

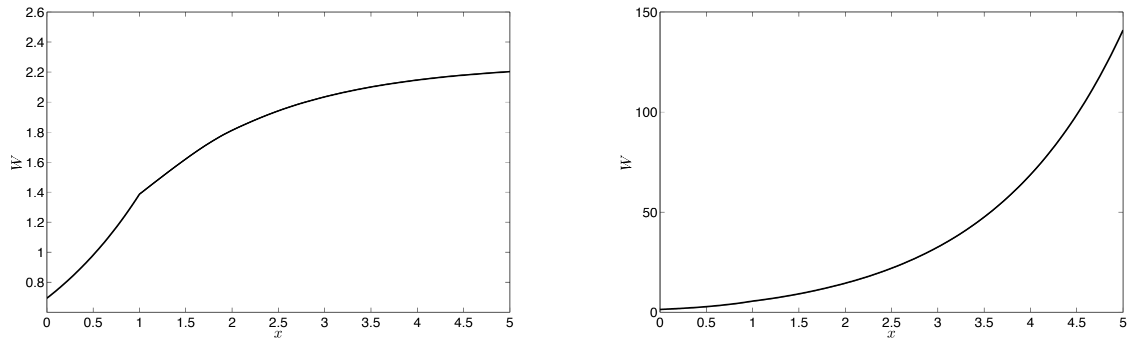
Therefore, the r.v. $S_\nu = \sum_{i=0}^{\nu} U_i$ has a Dirac mass at 0 with probability $1 - (1/\beta)$, and a density distribution $p_u^{*\nu}(x) = (1 - (1/\beta)) \sum_{j=1}^{\infty} \beta^{-j} p_u^{*j}(x)$.
(ii) Integrating $p_u^{*\nu}$ allows one to derive the cumulative distribution and thus the scale function W (for $\beta > 1$) as follows:

$$W(x) = \frac{1}{\beta} \sum_{k=0}^{\lfloor x \rfloor} \frac{(-1)^k}{k!} \big((x-k)/\beta\big)^k \exp\big((x-k)/\beta\big). \tag{51}$$

Similarly, upon elementary calculations and noting that $\frac{(-1)^k}{k!} \int_0^x y^k e^y dy = e^x \sum_{j=0}^{k} \frac{(-x)^j}{j!} - 1$, we can derive the closed form of the primitive $\widehat{W}(x) = \int_0^x W(y)dy$ as follows:

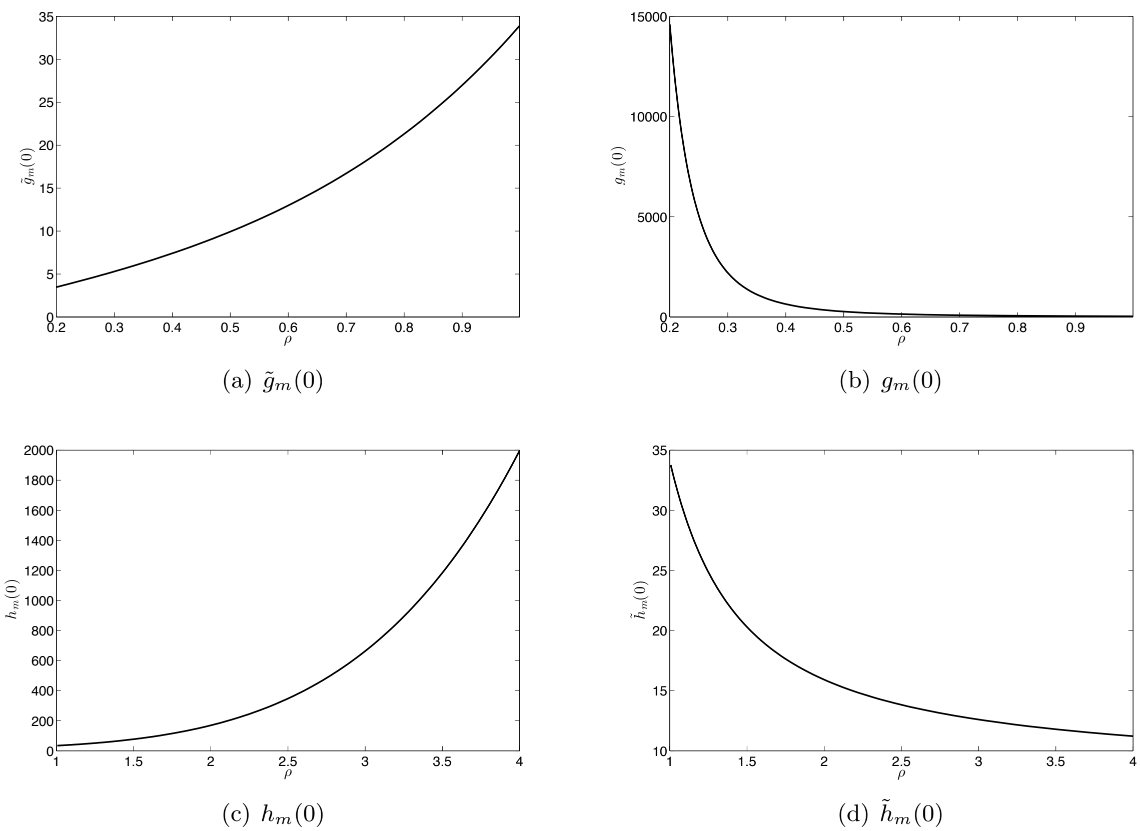
$$\widehat{W}(x) = \sum_{k=0}^{\lfloor x \rfloor} \Big(e^{(x-k)/\beta} \Big(\sum_{i=0}^{k} \frac{(-1)^j}{j!} \big((x-k)/\beta\big)^j \Big) - 1 \Big). \tag{52}$$

In Figure 4, we depict the scale function as a function of x for $\beta > 1$ and $\beta < 1$. We fix the parameters β respectively equal to 1.5 and 0.5. When $\beta < 1$, one should permute the role of W and $\widetilde{W}$ in order to use Equations (51) and (52) as W is not anymore a cumulative distribution function. Thus, to compute the scale function W, we first compute $\widetilde{W}$ using (52) to characterize $\widetilde{\mathbb{P}}(S_\nu \leq x)$ with $\beta(\rho)/\rho$ and then write $W(x) = \rho^{-x}\widetilde{W}(x)$.
(iii) In Figure 5, for a fixed threshold level $m = 5.5$, we represent the performance of the cusum procedure for different values of ρ both for an increase and a decrease of the intensity. This figure was depicted using the series representation of W and $\widehat{W}$ and the performance functions closed forms in Equations (38), (39) and (36). This represents the average delay until detection as well as the false alarm constraint. We observe that the detection is quicker as ρ moves off the critical value 1. Moreover, as noted in Section 5, we can see that the functions h_m and $\tilde{g}_m$ have similar behavior. This is also the case for $\tilde{h}_m$ and g_m. However, when ρ increases, we observe an instability in the numerical calculation of $\tilde{h}$. This phenomenon is not present for g_m, see Figure 6, even for large ρ.



(a) $\rho = 2$ (b) $\rho = 0.5$

FIGURE 4. Scale function $W(x)$ for different values of ρ.

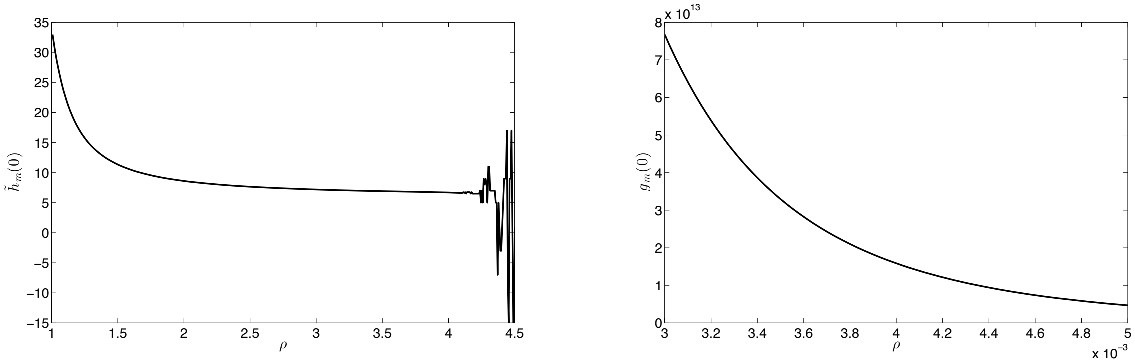


(a) $\tilde{g}_m(0)$ (b) $g_m(0)$

(c) $h_m(0)$ (d) $\tilde{h}_m(0)$

FIGURE 5. Performance of the cusum as a function of the parameter ρ with $m = 5.5$.

7.2. *Numerical issues.* (i) For $\rho > 1$, when either m or ρ becomes too large, some numerical instability may arise in the numerical calculation of $\tilde{h}_m(0)$ (Figures 6(a) and 7). This is mainly due to the particular series representation of the scale function W, where we are summing alternating individual terms that increase fast in absolute value. The phenomenon has also been observed for



(a) $\tilde{h}_m(0)$ (b) $g_m(0)$

FIGURE 6. Numerical (in-)stability in the calculation of $\tilde{h}_m(0)$ and $g_m(0)$ as a function of ρ for $m = 5.5$.

a long time in sequential hypotheses tests, DeLucia and Poor [12], or in ruin theory by Picard and Lefèvre [27] or Rullière and Loisel [32], where the scale function W plays also a central role. In any case, the problem is to find a well-conditioned algorithm to solve the delayed equation.

(ii) In DeLucia and Poor [12], a synthesis and fine analysis of the problem are proposed and

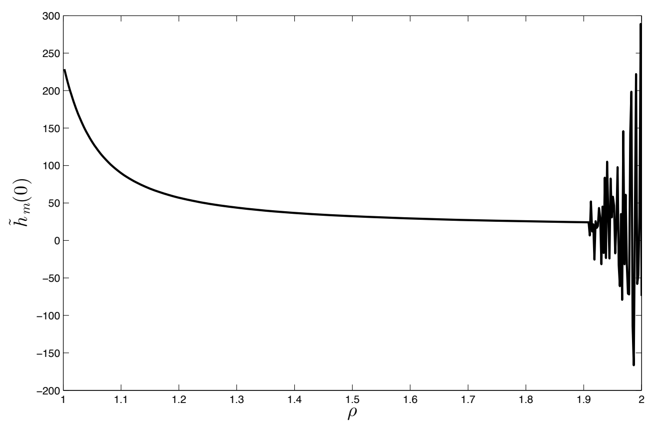


FIGURE 7. $\tilde{h}_m(0)$ as a function of the post-change parameter ρ for $m = 15$.

different solutions are tested, for example a solution based on the inversion of the Laplace transform given in Corollary 1. As tested in [12], the best approach is based on a recursive solution: "*a method that may be useful in solving other than the first-order DDE's*". The idea is to solve the problem recursively on the discrete intervals, $[0, 1), [1, 2)....[k, k+1)$ using that on $[0, 1)$, $W(x) = (1/\beta)e^{x/\beta}$, and on $[k, k+1)$, $Q_k(\xi) = e^{-(k+\xi)/\beta}W(k+\xi))$, $\xi \in (0, 1)$. It is clear that Q_0 is a constant function and Q_k is a polynomial of degree k. On the other hand, the a.e. continuity of W induces a continuity constraint on the polynomials Q_k at 0 and 1, i.e. $Q_k(1) = Q_{k+1}(0)$. Moreover, the DDE (21) implies a hierarchical relation on the polynomials derivatives $Q'_k(\xi) = -\alpha Q_{k-1}(\xi)$ with $\alpha = (1/\beta)e^{1/\beta}$. This recursion defines a set of polynomials that falls into the definition of Appell polynomials, introduced by Picard and Lefèvre [27] in ruin theory. Applying Taylor's formula to $Q_k(\xi)$, we obtain the

expansion $Q_k(\xi) = \sum_{i=0}^{k} Q_{k-i}(0)\frac{(\alpha\xi)^i}{i!}$ where the coefficients $Q_j(0)$ can be deduced recursively.
(iii) As pointed by DeLucia and Poor [12], this method is useful to avoid the unstable numerical outputs observed on the computation of W for large ρ. In Figure 8, we plot the function $\tilde{h}_m(0)$, for a range of ρ where the instability may arise with a fixed barrier level $m = 5.5$. We thus depict the function W obtained separately from the recursive (dashed line) as well as the series representation (solid line). The series form is numerically unstable, whereas the recursive solution remains accurate.

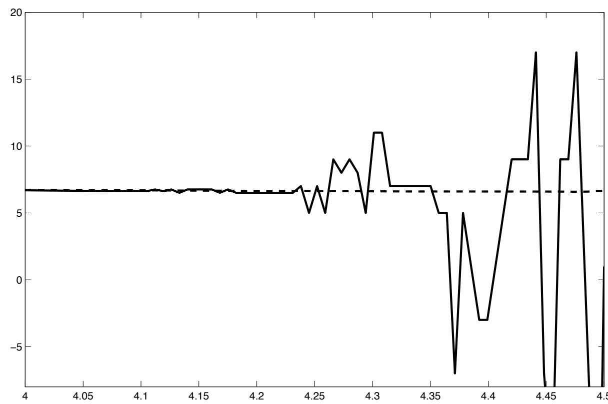


FIGURE 8. Series and recursive representation based computation of the performance function $\tilde{h}_m(0)$ for $m = 5.5$.

8. Conclusion.

Initially motivated by a problem of quickest detection of change in some longevity patterns, we have considered and solved the exact optimality of the minimax robust detection of a disorder time in the Poisson rate, with a self-contained presentation. In this Lorden-type context, the cusum stopping rule is shown to be optimal both for an increase or a decrease in intensity after the change.
Given the abundant literature on sequential testing and quickest detection, it may be surprising that this classical problem has not been solved earlier. We believe that this is due to the difference between the cases of increase and decrease, the former featuring non-continuous performance functions and requiring the use of a discontinuous local time.
As scale functions appear in the proof, one may wonder if it is possible to extend this result to a broader class of Lévy processes. This is left for further research, as well as some detailed analysis of the adaptation of this detection strategy for different sets of insurance data and the comparison with other detection strategies.

Acknowledgment. This work benefited from the financial support of the ANR project "LoLitA" (ANR-13-BS01-0011), Fondation du Risque Chair "Risques Financiers", BNP Paribas Cardif Chair "Management de la Modélisation" and Milliman Paris Chair "Actuariat Durable".

References.

[1] L.B. Afonso, R.M.R. Cardoso, and A.D.E. Dos Reis. Dividend problems in the dual risk model. *Insurance: Mathematics and Economics*, 53(3):906–918, 2013. 7, 11
[2] S. Asmussen. *Applied probability and queues.* Springer, 2003. 7, 8

[3] S. Asmussen and H. Albrecher. *Ruin probabilities*, volume 14. World Scientific Publ. Cy., 2010. 12

[4] B. Avanzi, H.U. Gerber, and E.S.W. Shiu. Optimal dividends in the dual model. *Insurance: Mathematics and Economics*, 41(1):111–123, 2007. 7, 11

[5] P. Barrieu, H. Bensusan, N. El Karoui, C. Hillairet, S. Loisel, C. Ravanelli, and Y. Salhi. Understanding, modeling and managing longevity risk: Key issues and main challenges. *Scandinavian Actuarial Journal*, 2012: 203–231, 2012. 1

[6] M. Basseville and I.V. Nikiforov. *Detection of Abrupt Changes in Signals and Dynamics Systems*. Springer-Verlag, 1986. 1, 3, 22

[7] E. Bayraktar, S. Dayanik, and I. Karatzas. The standard Poisson disorder problem revisited. *Stochastic processes and their applications*, 115(9):1437–1450, 2005. 2

[8] M. Beibel. A note on Ritov's bayes approach to the minimax property of the cusum procedure. *The Annals of Statistics*, 24(4):1804–1812, 1996. 2

[9] J. Bertoin. *Lévy processes*, volume 121. Cambridge Univ. Press, 1998. 2, 5, 8, 9, 14

[10] D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society*, pages 187–220, 1972. 3

[11] B. De Finetti. Su un'impostazione alternativa della teoria collettiva del rischio. In *Transactions of the XVth International Congress of Actuaries*, volume 2, pages 433–443, 1957. 10

[12] J. DeLucia and H.V. Poor. Performance analysis of sequential tests between Poisson processes. *IEEE Transactions on Information Theory*, 43(1):221–238, 1997. 25

[13] A. Dvoretzky, J. Kiefer, and J. Wolfowitz. Sequential decision problems for processes with continuous time parameter. Testing hypotheses. *The Annals of Mathematical Statistics*, 24(2):254–264, 1953. 2

[14] J. Grandell. *Doubly stochastic Poisson processes*. Springer, 1976. 4

[15] T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer, 2009. 2

[16] J.O. Irwin. On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's type II. *Biometrika*, pages 225–239, 1927. 23

[17] A.E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer, 2006. 5, 9

[18] A.E. Kyprianou. *Gerber–Shiu risk theory*. Springer Science & Business Media, 2013. 4

[19] R.S. Lipster and A.N. Shiryaev. *Theory of martingales*. Kluwer, 1989. 5

[20] G. Lorden. Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, 42(6): 1897–1908, 1971. 1, 2, 3, 4, 18

[21] Y. Mei, S.W. Han, and K.-L. Tsui. Early detection of a change in Poisson rate after accounting for population size effects. *Statistica Sinica*, 21(2):597, 2011. 2, 3

[22] G.V. Moustakides. Optimal stopping times for detecting changes in distributions. *The Annals of Statistics*, 14 (4):1379–1387, 1986. 3

[23] G.V. Moustakides. Performance of CUSUM tests for detecting changes in continuous time processes. In *Proceedings of the IEEE International Symposium on Information Theory*, page 186. IEEE, 2002. 3

[24] G.V. Moustakides. Optimality of the CUSUM procedure in continuous time. *The Annals of Statistics*, 32(1): 302–315, 2004. 1, 2, 4, 18

[25] E.S. Page. Continuous inspection schemes. *Biometrika*, 41(1/2):100–115, 1954. 1, 2

[26] G. Peskir and A.N. Shiryaev. *Advances in Finance and Stochastics*, chapter Solving the Poisson disorder problem, pages 295–312. Springer, 2002. 8

[27] P. Picard and C. Lefèvre. The probability of ruin in finite time with discrete claim size distribution. *Scandinavian Actuarial Journal*, 1997(1):58–69, 1997. 25

[28] M.R. Pistorius. On exit and ergodicity of the spectrally one-sided Lévy process reflected at its infimum. *Journal of Theoretical Probability*, 17(1):183–220, 2004. 2, 8, 16

[29] H.V. Poor and O. Hadjiliadis. *Quickest detection*. Cambridge Univ. Press, 2009. 2

[30] N.U. Prabhu. *Stochastic Storage Processes: queues, insurance risk, and dams, and data communication*. Number 15. Springer, 1998. 8

[31] T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels. *Stochastic processes for insurance and finance*, volume 505. Wiley, 2009. 7

[32] D. Rullière and S. Loisel. Another look at the Picard-Lefèvre formula for finite-time ruin probabilities. *Insurance: Mathematics and Economics*, 35(2):187–203, 2004. 25

[33] A.N. Shiryaev. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications*, 8(1):22–46, 1963. 1

[34] A.N. Shiryaev. Minimax optimality of the method of cumulative sums (CUSUM) in the case of continuous time. *Russian Mathematical Surveys*, 51(4):750–751, 1996. 2, 3, 18

[35] A.N. Shiryaev. On stochastic models and optimal methods in the quickest detection problems. *Theory of

Probability & Its Applications, 53(3):385–401, 2009. 1, 2

[36] S. Zacks. Exact determination of the run length distribution of a one-sided CUSUM procedure applied on an ordinary Poisson process. *Sequential Analysis*, 23(2):159–178, 2004. 12

L.P.M.A.
Université Pierre et Marie Currie
Case courrier 188
75252 Paris Cedex 05
France
E-mail: elkaroui@gmail.com

I.S.F.A.
Université de Lyon,
Université Claude Bernard Lyon 1,
50 Avenue Tony Garnier, F-69007 Lyon
France
E-mail: stephane.loisel@univ-lyon1.fr
E-mail: yahia.salhi@univ-lyon1.fr